

Université de Sherbrooke

**IDENTIFICATION DES ELEMENTS *CIS* D'ARN ET DEVELOPPEMENT D'UN
GENE RAPPORTEUR POUR CARACTERISER LES FACTEURS D'EPISSAGE
QUI CONTROLANT L'EXPRESSION DU FACTEUR DE TRANSCRIPTION
PRO-APOPTOTIQUE TAF65**

Par

KAMTCHUENG Catherine

Département de Microbiologie et d'Infectiologie

Mémoire présenté à la Faculté de Médecine et des Sciences de la Santé
en vue de l'obtention du grade de Maître ès sciences (M. Sc.) en Microbiologie

Sherbrooke, Québec, Canada

Janvier 2013

Membres du jury d'évaluation

Brendan BELL PhD, Département de Microbiologie et d'Infectiologie

Benoît CHABOT PhD, Département de Microbiologie et d'Infectiologie

Martin BISAILLON PhD, Département de Biochimie



Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-95159-0

Our file Notre référence

ISBN: 978-0-494-95159-0

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

RESUME

IDENTIFICATION DES ELEMENTS *CIS* D'ARN ET DEVELOPPEMENT D'UN GENE RAPPORTEUR POUR CARACTERISER LES FACTEURS D'EPISSAGE QUI CONTROLONT L'EXPRESSION DU FACTEUR DE TRANSCRIPTION PRO-APOTOTIQUE TAF6 δ

Par

KAMTCHUENG Catherine

Département de Microbiologie et d'Infectiologie

Mémoire présenté à la Faculté de Médecine et des Sciences de la Santé
en vue de l'obtention du diplôme de Maître ès sciences (M.Sc.) en Microbiologie,
Faculté de Médecine et des Sciences de la Santé, Université de Sherbrooke, Sherbrooke,
Québec, Canada, J1H 5N4

L'apoptose est un processus primordial pour le développement et le maintien des organismes eucaryotes. Elle est régulée à différents niveaux de l'expression génique. En fonction des stimuli intra- et extra-cellulaires, les facteurs de transcription régulent l'expression des protéines pro-survies et pro-apoptotiques. Ces facteurs de transcription facilitent la formation du complexe de pré-initiation (CPI) de la transcription pour activer la transcription. Le CPI est composé de l'ARN polymérase II et des facteurs généraux de transcription dont le facteur de reconnaissance du promoteur, TFIID. TFIID est un complexe multi-protéique composé de la protéine de liaison de la boîte TATA (TBP) et de 14 facteurs associés à TBP (TAFs) tel que TAF6 qui nous intéresse. TAF6 est exprimé sous 5 isoformes d'épissage alternatif (α , β , γ , δ , ϵ). TAF6 α et TAF6 δ sont deux entités antagonistes. Ils sont issus de l'épissage alternatif du deuxième exon de TAF6. Contrairement au variant majoritaire TAF6 α qui a une activité oncogénique et anti-apoptotique, le variant minoritaire TAF6 δ est pro-apoptotique. A l'opposé de TAF6 α , l'excision de la partie alternative du deuxième exon empêche la dimérisation de TAF6 δ avec TAF9 dans TFIID (TFIID π).

Les analyses de puce à ADN ont montré que l'impact sur le transcriptome de la perte de TAF6 α est fortement distinct de l'impact causé par l'induction de l'isoforme pro-mort TAF6 δ par des oligonucléotides antisens qui bascule l'épissage (SSO pour Splice site Switching Oligonucleotids). En outre, la déplétion du variant d'épissage majeur TAF6 α

aboutit à la perte de la viabilité des cellules. L'importance de l'induction de TAF6 δ dans la mort cellulaire programmée et nos résultats précédents montrant que TAF δ induit l'apoptose indépendamment de p53 dans de nombreux types de cellules cancéreuses, nous ont incités à entreprendre une dissection des éléments *cis* d'ARN qui contrôlent l'épissage du variant TAF6 δ . Nous avons développé un système de minigène pour étudier les éléments *cis* d'ARN qui contrôlent l'expression de TAF6 δ . Le minigène inclut la séquence génomique de TAF6 (exon 2 à exon 3) qui est dirigé par le promoteur CMV et il mime le patron d'épissage alternatif de TAF6 α et TAF6 δ endogène. Nous avons entrepris une analyse mutationnelle pour identifier les éléments *cis* de l'ARN pré-messager de TAF6. Nos données ont mis en évidence un site activateur d'épissage dans l'exon 2 constitutif. Dans l'intron, deux motifs polyC et polyG pourraient réguler l'épissage alternatif de TAF6 δ . Ces motifs représentent des sites potentiels de liaison pour les protéines de liaison à l'ADN, hnRNP K et H respectivement. Nous avons donc testé l'effet de la surexpression de hnRNP K et H sur l'épissage alternatif de TAF6 et nous n'avons eu aucuns changement sur l'expression de TAF6 δ endogène. De plus, nous avons constaté que la mutation d'un seul nucléotide qui semble perturber la structure secondaire d'ARN au site d'épissage proximal, renverse complètement le patron d'épissage. Cette mutation favorise le choix du site d'épissage distal (un site faible) à la place du site d'épissage proximal (TAF6 α). Nos résultats suggèrent aussi qu'un élément régulateur d'épissage qui favorise TAF6 α est présent dans l'exon 2 alternatif.

Pour permettre l'identification des facteurs d'épissage influençant le choix du site d'épissage, nous avons créé un nouveau système d'épissage rapporteur. Notre nouveau vecteur contient la séquence génomique de TAF6 (exon 2 à exon 4) modifiée par l'introduction d'un codon stop prématuré (CSP) dans l'exon 2 alternatif et fusionné à la protéine EYFP. La combinaison des essais de transfections transitoires avec les SSOs ont été utilisés pour valider ce système contrôlé par cytométrie de flux. Dans le futur, ce système pourrait être utilisé pour produire une lignée stable afin d'identifier les facteurs d'épissage impliqués dans la régulation de l'épissage alternatif par un criblage d'inhibition (siRNA) ou de surexpression (ADNc).

Donc, mon projet présente la première identification des éléments *cis* qui contrôlent l'épissage du facteur aussi bien que le développement d'un système d'épissage rapporteur créé pour permettre l'identification des protéines d'épissage qui régulent l'expression de TAF6 δ .

Mots-clés : TAF6 δ , apoptose, épissage alternatif, éléments *cis*, facteurs d'épissage

SUMMARY

IDENTIFICATION DES ELEMENTS CIS D'ARN ET DEVELOPPEMENT D'UN GENE RAPPORTEUR POUR CARACTERISER LES FACTEURS D'EPISSAGE QUI CONTROLONT L'EXPRESSION DU FACTEUR DE TRANSCRIPTION PRO-APOTOTIQUE TAF65

By

KAMTCHUENG Catherine

Department of Microbiology and Infectiology

Master's thesis is presented to the Faculté de Médecine et des Sciences de la Santé
In order to obtain the diploma of Master of Sciences (M.Sc.) in Microbiology,
Faculté de Médecine et des Sciences de la Santé, Université de Sherbrooke, Sherbrooke,
Québec, Canada, J1H 5N4

Apoptosis is crucial to the development and homeostasis of eukaryotic organisms. It is regulated at different levels of gene expression. In response to intra- and extra-cellular stimuli, transcription factors regulate the expression of pro-survival and pro-apoptotic proteins. These transcription factors facilitate the formation of the RNA polymerase II pre-initiation complex (PIC) to activate transcription. The PIC is composed of RNA Polymerase II (Pol II) and the general Pol II transcriptional factors including the major core promoter recognition factor TFIID. TFIID is a multiprotein complex containing the TATA-Binding Protein (TBP) and 14 TBP Associated Factors (TAFs), including TAF6, which is the focus of my master's thesis. TAF6 possess 5 alternative splicing isoforms (α , β , γ , δ , ϵ). TAF6 α and TAF6 δ possess opposing biological activities. They are produced via the alternative splicing of exon2 TAF6. In contrast to the major alpha variant that possesses anti-apoptotic and oncogenic activities, the minor delta variant is pro-apoptotic. In contrast to TAF6 α , TAF6 δ doesn't dimerize with TAF9 in TFIID because it lacks the alternative exon2 that is excluded by an alternative splicing event.

We used microarray analysis to show that the impact on the transcriptome of the loss of the TAF6 α is highly distinct from the impact caused by induction of pro-death TAF6 δ isoform by antisense splice switching oligonucleotides. Furthermore, the depletion the major TAF6 α splice variant results in the loss of cell viability. The importance of the induction of TAF6 δ in programmed cell death, taken together with previous results showing that TAF6 δ induces p53-independent apoptosis in numerous cancer cell types, prompted us to undertake a dissection of the cis-acting RNA elements

that control splicing of the TAF6delta variant. We developed a minigene system to study *cis*-acting RNA elements that control the expression of TAF6delta. The minigene includes the genomic sequence of TAF6 (exon 2 to exon3) is driven by CMV promoter and it recapitulates the alternative splicing pattern of endogenous TAF6alpha and TAF6delta. We employed a mutational analysis to identify *cis*-acting elements of the TAF6 pre-mRNA. Our data have identified one enhancer splice element in the constitutive exon 2. Within intron 2, two polyC and polyG elements could regulate the alternative splicing of TAF6delta. These motifs represent potential binding sites for the RNA-binding proteins hnRNP K and H respectively. We therefore tested the effect of the overexpression of hnRNP K and H on TAF6 alternative splicing and found no effect on the expression of endogenous TAF6delta. Moreover, we report that a single nucleotide change that appears to perturb RNA secondary structure at the proximal splice site, completely reverses the splicing pattern. This mutation favors the choice of the distal splice site (weakened site) instead of the proximal splice site (TAF6alpha). Our data also suggest that a splice regulatory element that favors TAF6alpha expression is present in alternative exon 2.

To enable the identification the splicing factors influencing the choice of splicing site, we designed a new splicing reporter system. Our new vector contains the sequence of our minigene added a premature truncated codon in the alternative exon 2 and fused in frame to the EGFP protein. The combination of transient transfection assays with splice site switching oligonucleotides was used to validate this system as monitored by flow cytometry. In the future this system could be used to produce stable reporter cells lines to identify the splicing factors involved to the regulation of alternative splicing by screening cDNA or siRNA libraries.

Thus, my thesis presents the first identification of *cis*-acting RNA elements that control the splicing of the pro-apoptotic splicing factor TAF6delta as well as the development of a splice reporter system designed to enable the identification of splicing factors that regulate TAF6delta expression.

Keywords : TAF6delta, apoptosis, alternative splicing, *cis*-acting elements, splicing factors

A vous, ma chair,
Feu Jean-Emmanuel KAMTCHUENG,
Ivonne Brigitte MENEFIN,
Elisabeth Candice TUMMAMO,
Pauline Emmanuelle METAYOU,
Julienne Clara DJEMMO,

A mes grand-parents, Feu NEMOGNE - TUMMAMO, Feu NEKAM - METAYOU,
A Feu Victor OUAMBO et Catherine NGUEMKAM.

**« Le cœur de l'homme intelligent acquiert la connaissance,
et l'oreille des sages cherche la connaissance. »
Proverbes 18 : 15**

TABLE DES MATIERES

RESUME	II
TABLE DES MATIERES	VIII
LISTE DES TABLEAUX	XI
LISTE DES FIGURES	XII
LISTE DES ABBREVIATIONS	XIV
 INTRODUCTION	 1
 I. L'apoptose	 2
A. Définition	2
B. Morphologie	2
C. Mécanismes	3
1. La voie extrinsèque	4
1.1 Les caspases	4
1.2 Mécanisme	6
2. La voie intrinsèque	7
2.1 La famille Bcl-2	7
2.2 Mécanisme	9
II. La transcription de l'ADN par l'ARN polymérase II	10
A. Le promoteur de base	11
B. L'initiation de la transcription	13
1. Mécanisme	13
2. TFIID	15
C. L'élongation et la terminaison de la transcription	20
III. TAF6	21
A. TAF6	21
B. TAF68	22

IV. L'épissage et l'épissage alternatif de l'ARN	24
A. Les séquences qui définissent les introns	24
B. La réaction chimique de l'épissage	26
C. Le spliceosome	26
D. Le mécanisme général de l'épissage	27
E. L'épissage alternatif	28
1. Les séquences exoniques/introniques et les facteurs d'épissage	29
1.1 Les protéines SR	30
1.2 Les hnRNPs	32
2. La transcription couplée a l'épissage alternatif	33
3. L'influence des structures secondaires sur l'épissage alternatif	34
4. La régulation de l'épissage par les ARNs non codants	35
LES OBJECTIFS	36
MATERIELS ET METHODES	38
LES RESULTATS	54
I. Chapitre 1 : l'article	55
Avant- propos de l'article	55
Résumé de l'article	55
Abstract	56
Introduction	57
Materials and methods	60
Results	63
Discussion and conclusions	74
Acknowledgements	76
Authors contributions	77
Conflict of interest	77

References	77
Legends of figures	82
II. Chapitre 2	100
A. Les facteurs hnRNPK et hnRNPH interviennent-ils à la régulation de l'épissage alternatif de TAF6 ?	100
B. Développement d'un nouveau vecteur d'expression : pEYFP-TAF6-PTC	102
 DISCUSSION ET CONCLUSION	109
REMERCIEMENTS	115
LISTE DES REFERENCES	116

LISTE DES TABLEAUX

Tableau 1 : Les facteurs généraux de transcription	15
Tableau 2 : Les facteurs associés à TBP (TAFs)	19
Tableau 3 : Liste des séquences des oligonucléotides	40

LISTE DES FIGURES

Figure 1 : Les voies de la signalisation de l'apoptose	4
Figure 2 : Les caspases	6
Figure 3 : La famille Bcl-2	8
Figure 4 : Le promoteur de base	12
Figure 5 : Le facteur général de transcription, TFIID	16
Figure 6 : Le domaine de repliement de type d'histone des variants d'épissage de TAF6 et dimérisation de TAF6α avec TAF9	18
Figure 7 : Les isoformes de TAF6	22
Figure 8 : Le modèle de la voie de TAF6δ	23
Figure 9 : Les séquences consensus des sites d'épissage	25
Figure 10 : Le mécanisme d'épissage	28
Figure 11 : Les différents types d'épissage alternatif	29
Figure 12 : La régulation de l'épissage alternatif	30
Figure 13 : Les différentes classes des protéines SR	31
Figure 14 : Les différentes hnRNPs	32
Figure 1 de l'article : The role of alternative splicing in the TAF6δ pathway of apoptosis ...	88
Figure 2 de l'article : TAF6 is essential for human cell viability	89
Figure 3 de l'article : Distinct impact of TAF6δ induction versus total TAF6 mRNA depletion on the transcriptome of HeLa cells	90
Figure 4 de l'article : Development and validation of a minigene system to study the alternative splicing of TAF6	91
Figure 5 de l'article : Scanning mutagenesis of constitutive exon 2 and exon 3	92
Figure 6 de l'article : Targeted mutagenesis of intron 2 of the <i>taf6</i> gene	93
Figure 7 de l'article : RNA secondary structure at the proximal 5' splice site can force the selection of TAF6δ	94
Figure 8 de l'article : Evidence for an exonic splicing silencer in alternative exon 2	95
Figure 9 de l'article : A hypothetical model for TAF6δ alternative splicing	96
Figure S1 de l'article : Morphology of cell death in response to siRNA-mediated depletion of TAF6	97
Figure S2 de l'article : Mapping of the branchpoint in intron 2 of the TAF6 minigene	98
Figure S3 de l'article : Schematic representation of mutations used in this study	99
Figure 15 : La surexpression de hnRNPK et hnRNPH dans les cellules Hela ws	101
Figure 16 : Le profil d'épissage alternatif de TAF6 en fonction des quantités de hnRNPH et hnRNPK	102
Figure 17 : Le nouveau gène rapporteur EYFP-TAF6-PTC	104
Figure 18 : Les différentes mutations pour introduire un codon stop dans TAF6	104
Figure 19 : Les mutations dans l'exon2 alternatif pour introduire un codon stop	105
Figure 20 : Les plasmides dérivés de pEYFP-N1	106

Figure 20 : L'expression de pEYFP-TAF6-PTC dans les cellules Hela ws 107

LISTE DES ABBREVIATIONS

5'	Extrémité 5'
5'SS	Site d'épissage 5', 5' splicing site
3'	Extrémité 3'
3'SS	Site d'épissage 3', 3' splicing site
°C	Degré Celcius
A	Adénosine
ADN	Acide DésoxyriboNucléique
ADNc	Acide DésoxyriboNucléique complémentaire
ADP	Adénosine DiPhosphate
Apaf1	Apoptosis Protease Activating Factor 1
APS	Ammonium PerSulfate
ATP	Adénosine TriPhosphate
ARN	Acide RiboNucléique
ARNm	Acide RiboNucléique messenger
ARNmi	Acide RiboNucléique micro
ARNr	Acide RiboNucléique ribosomique
ARNsn	Acide RiboNucléique « small nuclear», petit Acide RiboNucléique
ARNsno	Acide RiboNucléique « small nucleolar », Petit acide ribonucléique
ARNt	Acide RiboNucléique de transfert
Asp	Aspartate
Bad	BCL2-associated agonist of cell death
Bak	BCL2-antagonist/killer 1
Bax	BCL2-associated X
Bcl 2	B-cell CLL/lymphoma 2
Bcl XL	B-cell lymphoma-extra large
BET	Bromure d'éthidium
BH	Bcl-2 Homology
Bid	BH3 Interacting Death Domain Agonist
Bik	BCL2-interacting killer

Bim	B-cell lymphoma 2 Interacting Mediator of cell death
Bmf	Bcl2 modifying factor
Bok	BCL2-related ovarian killer
BOP	BH3 Only Protein
BRE	TFIIB Responsive Element
CaCl ₂	Calcium di chloride, Dichlorure de calcium
CAD	Caspase activated DNase
CAK	Cdk Activating Kinase
CARD	Caspase Recruitment Domain
Caspase	CysteinyI Aspartate Specific ProteASE
CDK7	Cyclin-dependent kinase 7
CED9	CEll Death abnormality 9
c-FLIP	Cellular FLICE-inhibitory protein
Cl	Chlore
CO ₂	Dioxyde de carbone
CPSF	Cleavage and Polyadenylation Specific Factor
CS	Calf Serum
CStF	Cleavage Stimulation Factor
CTD	C-terminal Domain, Domaine C-terminal,
Cyt C	Cytochrome C
DCE	Downstream Core Element
DD	Death Domain
DED	Death Effector Domain
DISC	Death Inducing Signalling Complex
DMEM	Dubelcco's Modified Eagle's Medium
DMRIE-C	1,2-dimyristyloxy-propyl-3-dimethyl-hydroxy ethyl ammonium bromide- Cholestérol
dNTP	Désoxynucléotide triphosphate
DPE	Downstream Promoter Element
DTT	Dithiothréitol
EDTA	EthyleneDiamineTetraAcétic Acid, Acide éthylènediaminetétraacétique
ELL	Elongation factor RNA polymerase II

ERCC2	Excision Repair Cross-Complementing rodent repair deficiency, complementation group 2
ESE	Exonic Splicing Enhancer
ESS	Exonic Splicing Silencer
EYFP	Enhanced Yellow Fluorescent Protein
FADD	Fas Associated via Death Domain
FBS	Fetal Bovine Serum
g	G-force, RCF = Relative Centrifugal Force
GABAA α 2	Gamma-AminoButyric Acid receptor, subunit alpha 2
GTP	Guanosine TriPhosphate
H2A/H2B	Histone 2A/2B
H3/H4	Histone 3/4
HFD	Histone Fold Domain
hnRNP	Heterogenous nuclear RiboNucleoProtein
HRK	HaRaKiri, BCL2 interacting protein
IAP	Inhibitor of Apoptosis Protein
ICAD	Inhibitor of Caspase Activated DNase
Inr	Elément Initiateur
Intron	INTervening RegiON
ISE	Intonic Splicing Enhancer
ISS	Intronic Splicing Silencer
kb	Kilobase
KCl	Chlorure de potassium
KCM	Chlorure de potassium et de magnésium
kDa	Kilodalton
KH	hnRNPK Homology domain
KH ₂ PO ₄	Potassium Dihydrogen Phosphate
LB	Luria Broth
M	Molaire
MAC	Mitochondrial Apoptosis induced Channel
MAT 1	Methionine AdenosylTransferase I
Mcl-1	Myeloid Cell Leukemia sequence 1

MDM2	Murine Double Minute 2
MED17	MEDiator complex subunit 17
µg	Microgramme
MgCl ₂	Dichlorure de magnésium
µL	Microlitre
mL	Millilitre
µM	Micromolaire
mM	Millimolaire
MTE	Motif Ten Element
Na ₂ HPO ₄	Sodium phosphate dibasic, Disodium hydrogen phosphate
NaCl	Chlorure de sodium
NaOH	Hydroxyde de sodium
NER	Réparation de l'ADN par excision de nucléotides, « Nucleotide Excision Repair »
ng	Nanogramme
NLS	Nuclear localization Signal
Noxa	NADPH oxidase activator 1
nM	Nanomolaire
O ₂	Dioxygène
Opti-MEM	Opti- Minimal Essential Medium
pb	Paire de base
PBS	Phosphate Buffered Saline, Tampon salin de phosphate
PCD	Programmed Cell Death
PCR	Polymerase Chain Reaction, Réaction en chaîne par polymérase
pH	Potentiel hydrogène
PIC	Pre-Initiation Complex
PMAIP1	Phorbol-12-Myristate-13-acetate induced protein 1
Pol I	ARN polymérase I
Pol II	ARN polymérase II
Pol III	ARN polymérase III
PS	Phosphatidyl sérine
PTB	Polypyrimidine Tract Binding

PUMA	p53 Upregulated Modulator of Apoptosis
PWI	Motif de liaison alternatif de l'ARN
RGG	Domaine riche en glycine et arginine
ROCK I	Rho-associated coiled coil containing protein kinase 1
rpm	Rotations per minute
RRM	Motif qui reconnaît l'ARN, « RNA Recognition Motif »
RS	Domaine arginine et serine riche
RT	Reverse Transcription, transcription inverse
SDS	Sodium DodécylSulfate
SELEX	Systematic Evolution of Ligands by EXponential enrichment
SF3a/SF3b	Splicing factor 3a/3b
SSO	Splice switching oligonucleotide
SVF	Serum de veau foetal
Swi/Snf	SWItch/Sucrose NonFermentable
TAF	TBP Associated Factor
TBE	Tampon Tris-borate-EDTA
tBid	Truncated BH3 Interacting Death Domain Agonist
TBP	TATA Binding Protein
TE	Tris-EDTA
TEMED	Tetramethylethylenediamine
TFIID	Transcription factor IID
TFTC	Tata binding protein Free TAF containing Complex
TNF α	Tumor necrosis factor alpha
TNF-R1	Tumor necrosis factor receptor 1
U	Unité
U2AF35/65	U2 small nuclear RNA auxiliary factor 35/65
UsnRNP	Uracile rich Small nuclear ribonucleoprotein
UV	Ultra-violet
V	Volt
XPD	Xeroderma pigmentosum D
ZnM	Motif de doigts de zinc

INTRODUCTION

I. L' APOPTOSE

A. Définition

Le terme « apoptose » a été utilisé pour la première fois en 1972 par Kerr pour décrire la condensation des « corps apoptotiques » qui est une des caractéristiques de l'apoptose (Kerr *et al.*, 1972). Le nom « apoptose » fait référence à la chute programmée des feuilles d'automne. L'apoptose, aussi appelée « mort cellulaire programmée » (programmed cell death ou PCD) est un processus essentiel où la cellule intervient dans sa propre destinée, son suicide. Ce phénomène physiologique est impliqué dans le développement comme l'embryogénèse (la sculpture des mains, pieds) et dans l'homéostasie cellulaire (balance entre la prolifération et la différenciation) par exemple du système immunitaire où l'apoptose est responsable de la délétion des cellules T autoréactives (la tolérance du soi) et la sélection des lymphocytes B. Le dysfonctionnement de l'apoptose peut mener à diverses pathologies, soit par défaut pour l'oncogénèse et les maladies auto-immunes (diabète de type I, leucoencéphalie dégénérative) (Thompson, 1995 ; Connell *et al.*, 2008), soit par son hyper-activation favorisant le développement de pathologies neurodégénératives (les maladies de Parkinson, d'Alzheimer et de Charcot) (Singh et Dikshit, 2007 ; Cenini *et al.*, 2008).

B. Morphologie

Lorsque les cellules entrent en apoptose, elles subissent des changements morphologiques qui vont les isoler des autres cellules. On observe un rétrécissement cellulaire, un cytoplasme dense, des organelles plus denses et une condensation chromatiniennne suivie de la fragmentation du noyau. Les cellules se désintègrent en plusieurs corps apoptotiques. L'asymétrie membranaire des cellules est aussi modifiée au cours de l'apoptose. En effet les phosphatidyl-sérines (PS) de la couche interne de la membrane plasmique sont transloquées vers la couche externe. Les PS deviennent ainsi des signaux

phagocytaires pour les cellules avoisinantes concernées qui éliminent les cellules apoptotiques sans l'induction de réponse inflammatoire. La fragmentation de l'ADN a aussi lieu. Lors de l'activation de la cascade des caspases, la caspase 3 active clive ICAD, l'inhibiteur de l'endonuclease CAD (pour Caspase-activated DNase). CAD libéré migre vers le noyau pour fragmenter l'ADN chromosomique en plusieurs petits brins de 180pb (Strasser *et al.*, 2000).

C. Mécanismes

L'apoptose est un mécanisme dynamique qui nécessite l'activation de voies de signalisation intracellulaire. On distingue deux voies principales d'activation de l'apoptose (*Figure 1*) :

- La voie extrinsèque, activée par un signal extracellulaire dit « signal de mort ». Cette activation dépend de la liaison d'un ligand à son récepteur comme TNF α /TNF-R1, FasL/Fas, TRAIL/DR4 ou DR5 (Polunovsky *et al.*, 1994 ; Suda et Nagata, 1994 ; Kayagaki, 1999).
- La voie intrinsèque, activée par un signal intracellulaire comme les dommages à l'ADN par les UV, des médicaments, etc. (Kastan *et al.*, 1991 ; Zhang *et al.*, 1998). Elle met en jeu le facteur de transcription p53 et la mitochondrie (Chipuk et Green, 2006).

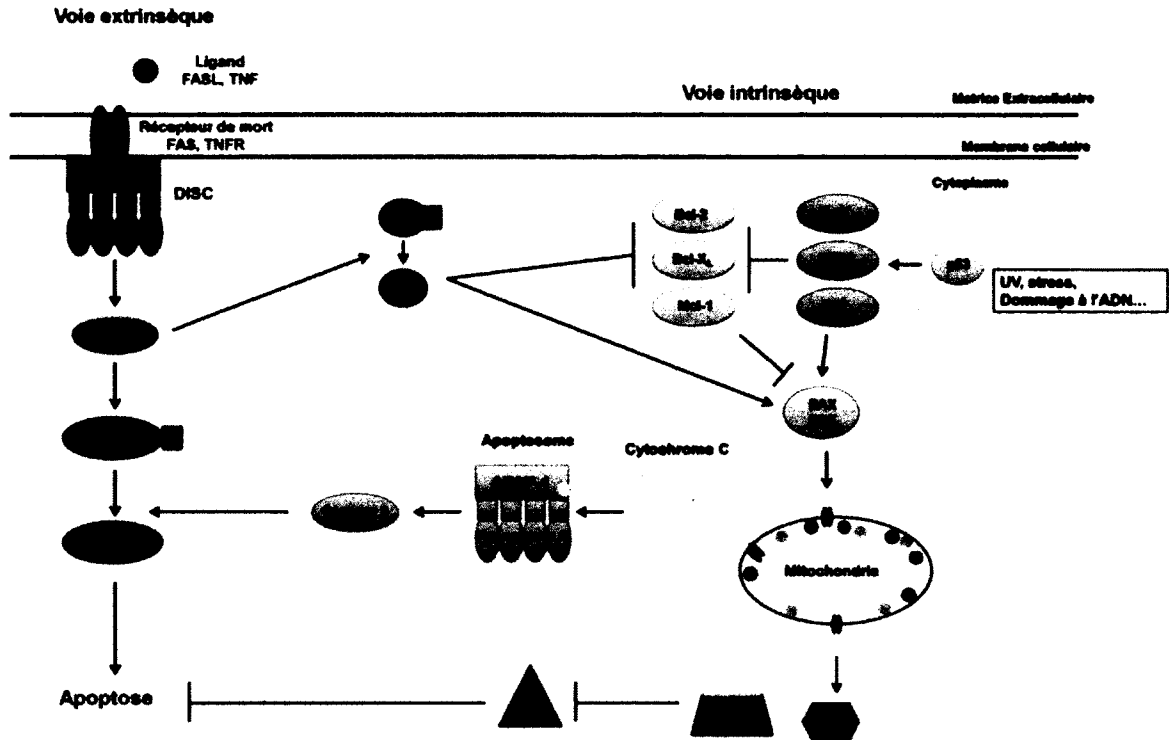


Figure 1 : Les voies de signalisation de l'apoptose

L'apoptose est induite par les voies intra- ou extra-cellulaires activées par leurs signaux distincts (adapté de Ashkenazi, 2008).

1. La voie extrinsèque

1.1. Les caspases

Les caspases (pour Cysteiny Aspartate Specific ProteASES) sont une famille de cystéine protéases dont le site actif comprend une cystéine dans une séquence conservée QACXG (Degterev *et al.*, 2003). Les caspases clivent des protéines inactives après un résidu d'acide aspartique d'une séquence consensus (D/L/V/W) E (H/T/V/X) D. Elles sont importantes pour l'induction et l'exécution de l'apoptose dans les deux voies. Les caspases sont synthétisées sous forme de pro-caspases (zymogènes) inactives. Leur

activation se fait par clivage au niveau du pro-domaine par d'autres caspases ou par elles mêmes lorsqu'elles sont dans un complexe d'induction d'apoptose.

Les caspases sont classées en trois groupes :

- Les caspases initiatrices qui possèdent un long pro-domaine: les caspases 2, 8, 9 et 10 activent d'autres caspases en clivant les pro-domaines.
- Les caspases effectrices qui possèdent un court pro-domaine: les caspases 3, 6 et 7 mènent à la mort cellulaire. Elles clivent aussi des protéines intervenant dans l'intégrité structurale du noyau et de la cellule (actine, lamine), des kinases (ROCKI), des inhibiteurs de nucléases (ICAD) etc.
- Les autres caspases : les caspases 1, 4, 5, 11, 12, 13, 14 participent à la régulation de la réponse inflammatoire, à la maturation des cytokines.

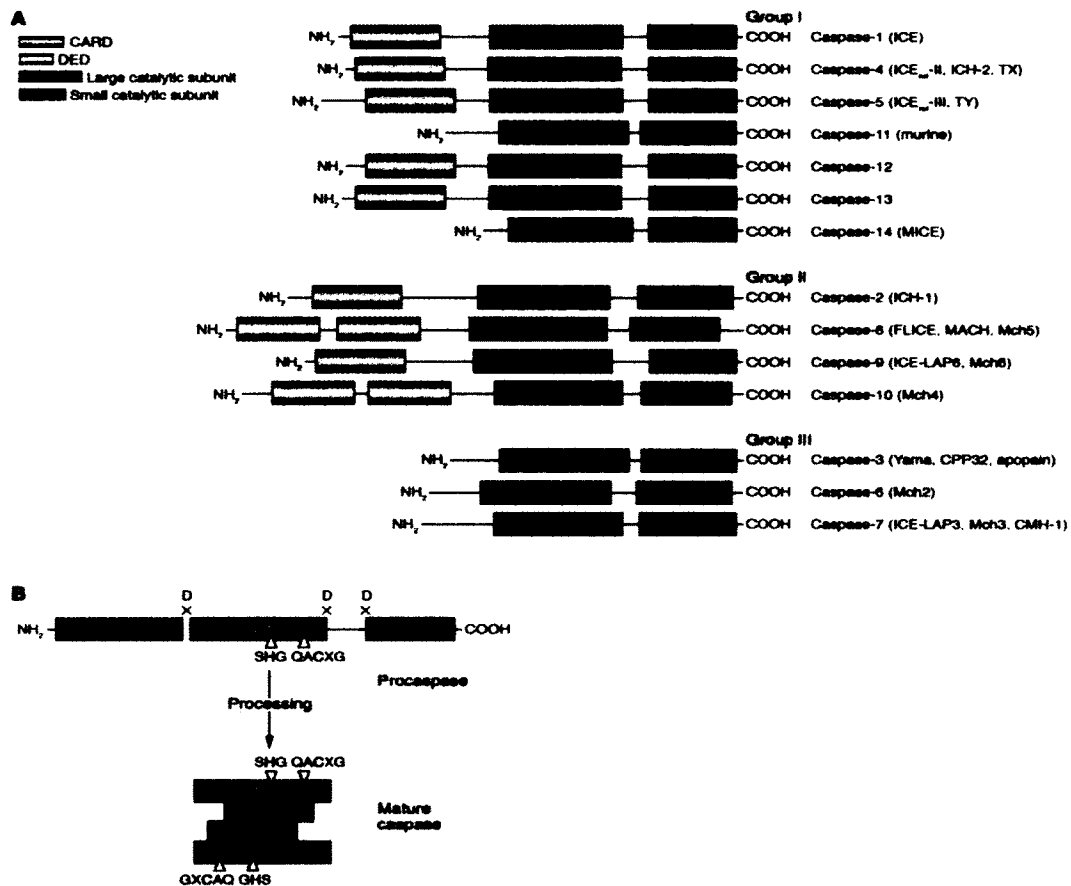


Figure 2 : Les caspases

A) Les caspases sont subdivisées en trois groupes : groupe I: les caspases inflammatoires ; groupe II: les caspases initiateuses de l'apoptose; groupe III: les caspases effectrices de l'apoptose. Le CARD, le DED et les sous-unités catalytiques, grande et petite, sont indiquées. CARD : CAspase ReCruiTment Domain ; DED : DeaTh Effector Domain. (B) Schéma de l'activation de la pro-caspase. Le clivage de la pro-caspase au site Asp-X spécifique mène à la formation d'une caspase mature par la constitution de l'hétérotétramère p20₂-p10₂ et la perte du pro-domaine. Les résidus impliqués dans la formation du centre actif sont montrés (tiré de Lavrik *et al.*, 2005).

1.2. Mécanisme

La voie extrinsèque est déclenchée par la liaison d'un ligand de mort extracellulaire à son récepteur de mort à la surface de la cellule sous forme de trimère. Les récepteurs activés

possédant des domaines intracellulaires DD (pour Death Domain), interagissent avec la protéine adaptatrice FADD (pour Fas-Associated via Death Domain) qui a aussi un domaine DD (Kischkel *et al.*, 1995). Le domaine DED (pour Death Effector Domain) de FADD permet le recrutement des pro-caspases initiateuses 8 et 10 qui possèdent un motif homologue. Ce complexe hétérotrimérique est appelé DISC (pour Death-Inducing Signalling Complex) (Kischkel *et al.*, 2001). Le recrutement des pro-caspases au sein du DISC provoque leur auto-clivage qui permet leur dimérisation et leur activation (Muzio *et al.*, 1998). Les caspases clivent ensuite les pro-caspases effectrices 3, 6 et 7 qui sont déjà sous forme de dimères dans le cytoplasme (Stennicke *et al.*, 1998). Les caspases effectrices vont par la suite cliver les protéines qui participent aux changements biochimiques et morphologiques des cellules apoptotiques.

Dans les cellules de type I (p. ex. les thymocytes), la caspase 8 peut activer directement la caspase 3. Par contre dans les cellules de type II (p. ex. les hépatocytes), l'activation nécessite le clivage de la protéine Bid (pour BH3 interacting death domain agonist) en tBid (pour truncated Bid) par la caspase 8 (Li *et al.*, 1998). Cette étape va amplifier le signal d'apoptose en passant par la mitochondrie et ainsi activer Bax et Bak.

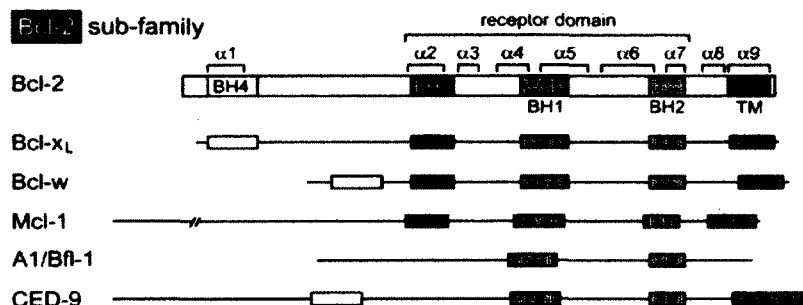
2. La voie intrinsèque

2.1. La famille Bcl-2

Cette famille de protéines est composée de trois sous familles dont les membres possèdent au moins un domaine BH (pour Bcl-2 homology, BH1 à BH4) (Cory et Adams, 2002). Le premier groupe est dit pro-survie ou anti-apoptotique (Bcl-2, Bcl-XL, Bcl-w, Mcl-1 et A1/Bfl1) dont les protéines contiennent trois à quatre domaines BH. Elles dimérisent avec les protéines de la deuxième sous famille pour empêcher l'induction de l'apoptose (Yin *et al.*, 1994). En effet, cette seconde sous famille (Bax, Bak, Bok), possédant trois domaines BH, a une activité pro-apoptotique (Hsu et Hsueh,

1998). Enfin la troisième sous famille (Bad, Bim, Bmf, Bid, Bik, HRK, EGL-1, Puma et Noxa) est constituée de protéines ne contenant que le domaine BH3 d'où leur nom de protéines BH3 isolé ou « BH3-only protein, BOP » ou les sensibilisateurs ou les activateurs (*Figure2*) (Subramanian *et al.*, 2007).

Pro-survival



Pro-apoptosis

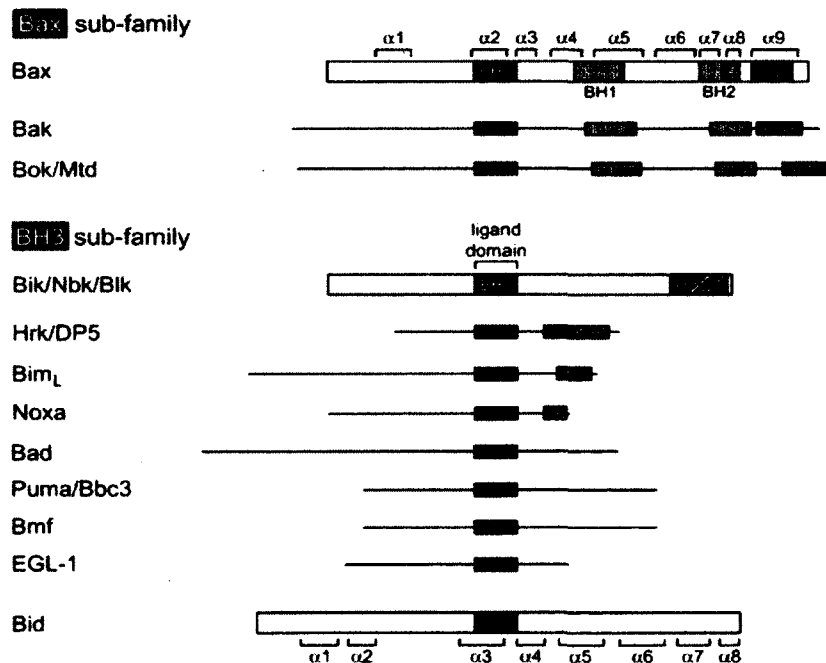


Figure 3 : La famille Bcl-2

La famille Bcl-2 est subdivisée en trois sous familles : 1) les pro-survies ou anti-apoptotiques (trois à quatre domaines BH) ; 2) les pro-apoptotiques (trois domaines BH) ; 3) les activateurs (domaine BH3 unique) (tiré de Cory S *et al.*, 2003).

2.2. Mécanisme

Les protéines BH3 isolés détectent les signaux d'apoptose par leur modification post-traductionnelle (Bad, Bid, Bim, Bmf) comme le clivage de Bid par la caspase 8 ou par leur induction transcriptionnelle (Noxa, Puma, HRK) généralement par le suppresseur de tumeur, p53. En réponse aux stress génotoxiques, p53 subit des modifications post-traductionnelles (par exemple la phosphorylation de la sérine 15) qui empêche son inhibition par MDM2 (pour Murine double minute 2) (Buschmann *et al.*, 2001). P53 devient stable et s'accumule dans le noyau. P53 peut ainsi contrôler l'expression de ses gènes cibles tels que Noxa/PMAIP1 (pour phorbol-12-myristate-13-acetate induced protein-1) et Puma (pour p53 upregulated modulator of apoptosis ou bbc3) (Kuribayashi et Deiry, 2008).

Ces activateurs permettent l'hétérodimérisation de Bax et de Bak. En effet l'activation de Bax par tBid lui permet de s'insérer dans la membrane mitochondriale externe. Par ailleurs l'induction de Noxa et Bad séquestre Mcl-1 (pour myeloid cell leukemia protein-1) et Bcl-XL respectivement pour libérer Bak pour son insertion dans la membrane mitochondriale externe. L'insertion des deux protéines pro-apoptotiques forme les pores MAC (pour Mitochondrial Apoptosis-induced Channel) à la surface externe de la mitochondrie. La perméabilisation de la membrane mitochondriale libère la protéine intermembranaire, le cytochrome c (cyt c) dans le cytoplasme. En présence d'ATP ou d'ADP, le cytochrome c interagit avec Apaf-1 (pour Apoptosis Protease Activating Factor-1). Grâce à la présence du domaine CARD (pour CAspase Recruitment Domain), Apaf-1 change de conformation et s'oligomériser avec d'autres Apaf-1 et la pro-caspase 9 qui est pourvu également d'un domaine CARD. Le multicomplexe composé de cytochromes c, d'Apaf-1 et de procaspases 9 est l'apoptosome au sein duquel la caspase 9 est activée par auto-clivage (Zou *et al.*, 1999). Enfin, la caspase 9 active les caspases effectrices 3, 6 et 7 aussi par clivage (Riedl *et al.*, 2001).

Au cours de l'activation de l'apoptose, la mitochondrie libère d'autres protéines pro-apoptotiques Smac/DIABLO qui lèvent l'inhibition des caspases effectrices par les IAPs

(pour Inhibitor of Apoptosis Protein), protectrices d'une activation anormale des caspases.

La mort cellulaire programmée est donc très importante pour le métabolisme cellulaire, finement régulée par la voie extrinsèque par l'intermédiaire du DISC formé après activation des récepteurs de mort ou bloqué par c-FLIP et des caspases activées par clivage ou inhibées par les IAPs. L'induction de l'apoptose, passant par la mitochondrie, dépend de la balance entre les protéines pro- et anti-apoptotiques qui sont régulées transcriptionnellement (p53), post-traductionnellement (phosphorylation, acétylation), et par épissage alternatif (Bcl-XL/XS, TAF6 α/δ). Afin de mettre plus en contexte notre thématique, les parties suivantes introduiront le facteur de transcription, TAF6 qui participe à l'initiation de la transcription et, par une isoforme d'épissage alternatif à l'induction de l'apoptose.

II. LA TRANSCRIPTION DE L'ADN PAR L'ARN POLYMERASE II

Chez les procaryotes, une seule ARN polymérase participe à la transcription pour tous les types d'ARN. Par contre chez les eucaryotes, il existe trois ARN polymérases :

- L'ARN polymérase I (Pol I) effectue la transcription des ARN ribosomiques (ARNr) 28S, 18S et 5,8S qui sont les plus représentés dans la cellule. Elle a lieu dans le nucléole à partir des gènes ribosomiques regroupés en « cluster ». La transcription produit un pré-ARNr de 47S qui subit des modifications post-transcriptionnelles pour donner les ARNr matures.
- L'ARN polymérase II (Pol II) catalyse la synthèse des ARN messagers (ARNm), codant pour les protéines. Elle synthétise aussi des microARN (ARNmi), certains petits ARN nucléaires (ARNsn pour small nuclear RNA) participant à l'épissage (à l'exception de l'ARN U6) et l'ARN de la télomérase.

- L'ARN polymérase III (Pol III) intervient dans la transcription des ARN de transfert (ARNt), de l'ARN ribosomique 5S ainsi que des petits ARN nucléolaires (snoRNA pour small nucleolar RNA) et de l'ARNsn U6.

Dans cette introduction, nous nous concentrerons sur la transcription des ARN messagers synthétisés par l'ARN polymérase II. La transcription est divisée en trois grandes étapes qui sont l'initiation (mise en place de la machinerie et synthèse des premiers nucléotides), l'élongation (allongement de la chaîne d'ARN) et la terminaison (libération du pré-ARN messenger).

A. Promoteur de base

L'initiation de la transcription est l'étape clé pour le choix des gènes à transcrire au cours de différents processus cellulaires. L'initiation de la transcription nécessite la reconnaissance du promoteur par plusieurs facteurs protéiques de transcription, permettant de réguler finement l'expression de ces gènes (Faiger *et al.*, 2007).

Chez les eucaryotes supérieurs, le promoteur de base est la combinaison de plusieurs éléments séquentiels suffisant pour l'initiation de la transcription (*Figure 4*) (Juven-Gershon et Kadonaga, 2010). La boîte TATA, dont la séquence consensus est TATA(A/T)A(A/T), est l'élément central du promoteur (Breathnach et Chambon, 1981). Elle est localisée à une trentaine de nucléotides en amont du site d'initiation (point +1). La fixation du facteur général de transcription TFIID via la TBP (TATA Binding Protein) à la boîte TATA jouerait un rôle dans le sens d'orientation de la transcription. L'élément Initiateur (Inr) a un rôle similaire à la boîte TATA (Smale et Baltimore, 1989). L'Inr se situe aux alentours du site d'initiation. Sa séquence consensus est Py-Py(C)-A⁽⁺¹⁾-N-T/A-Py-Py (Corden *et al.*, 1980). Il peut initier la transcription en synergie avec la boîte TATA (distance de séparation entre 25 et 30 pb) ou en son absence (O'shea-Greenfield et Smale, 1992). En effet la majorité des promoteurs sont dépourvus de la boîte TATA. Cet élément est reconnu par des sous-unités de TFIID (Bellorini *et al.*, 1996), les TAFs (TBP

associated Factors) TAF1 et TAF2 (Chalkley et Verrijzer, 1999). Les séquences BRE^u (Lagrange *et al.*, 1998) et BRE^d (Deng et Roberts, 2005) (pour TFIIB Responsive Element) sont situées en amont et en aval de la boîte TATA respectivement. Leurs séquences, GC riches, sont (G/C)-(G/C)-(G/A)-C-G-C-C. Ces deux séquences sont reconnues par un autre facteur général de transcription TFIIB. La séquence DPE (pour Downstream Promoter Element) est nécessaire à la liaison des promoteurs sans boîte TATA (Burke et Kadonaga, 1996). La séquence (A/G)-G-(A/T)-C-G-T-C est localisée à une trentaine de nucléotides en aval du site +1 d'initiation (Kutach et Kadonaga, 2000). Elle est reconnue par les TAF6 et TAF9 (Burke et Kadonaga, 1997). Le MTE (pour Motif Ten Element) est à 18 nucléotides en aval du site +1. Pour initier la transcription, il a besoin de l'Inr mais pas forcément de la boîte TATA et du DPE (Lim *et al.*, 2004). Enfin le DCE (pour Downstream Core Element) est une séquence tripartite qui s'étend de +6 à +34 en aval du site d'initiation. Il fonctionne avec la boîte TATA et l'Inr et il est reconnu par TAF1, une des sous unités de TFIID (Lewis *et al.*, 2000).

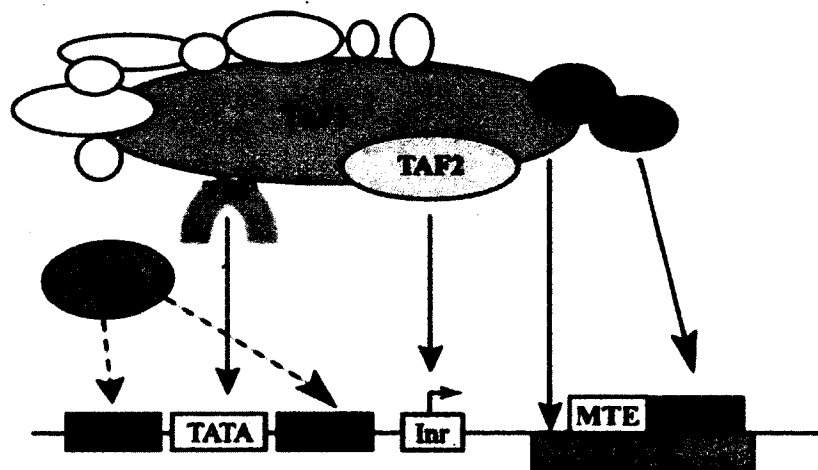


Figure 4 : Le promoteur de base

Représentation schématique des éléments constituant le promoteur de base et les facteurs de transcription s'y liant (tiré de Thomas et Chiang, 2006).

B. Initiation de la transcription

1. Mécanisme

L'initiation de la transcription nécessite l'intervention de complexes multiprotéiques, les facteurs généraux de transcription (TFIIA, -B, -D, -E, -F et -H) et de l'ARN polymérase II formant le complexe de pré-initiation (PIC) (*Tableau 1*) (Roeder, 1996).

La première étape de l'initiation de la transcription est la reconnaissance du promoteur par TFIID par l'intermédiaire TBP qui induit une courbure de l'ADN d'environ 90° (Roeder, 1996). Cette courbure permet le rapprochement des séquences régulatrices situées en aval et en amont du point +1 ainsi que l'ouverture de l'ADN. Lorsque le promoteur est pourvu de la boîte TATA, c'est la protéine TBP qui va s'y fixer. Par contre avec un promoteur dépourvu de la boîte TATA, ce sont les TAFs qui vont consolider la liaison de TFIID à l'ADN. Il existe un autre complexe TFIIIC (pour Tata binding protein Free TAF-containing Complex) ne possédant pas TBP, qui a les mêmes fonctions que TFIID (Muller et Tora, 2004). Donc TAF1 se lie à l'Inr et au DCE, TAF2 à l'Inr et TAF6/TAF9 au DPE.

TFIIA rejoint TFIID en liant TBP. TFIIA permet de stabiliser le complexe TFIID-ADN et ainsi de favoriser l'assemblage du complexe de pré-initiation (Buratowski *et al.*, 1989, Cortes *et al.*, 1992). TFIIB reconnaît TBP (Nikolov *et al.*, 1995) et les séquences BRE du promoteur de base (Reese, 2003). La liaison de TFIIB sur le BRE joue un rôle dans la détermination directionnelle (5'-3') de l'assemblage du PIC et du démarrage de la transcription (Lagrange *et al.*, 1998). Par la suite, TFIIB est reconnu par TFIIF, le « véhicule » de l'ARN polymérase II (Buratowski *et al.*, 1989). En effet, TFIIF lie l'ARN polymérase II pour supprimer les interactions non spécifiques à l'ADN. TFIIF aide la Pol II à se positionner au site +1 et stabilise le complexe de pré-initiation (Killeen et Greenblatt, 1992). La présence de TFIIF modifie à nouveau la courbure de l'ADN qui s'enroule autour de l'ARN polymérase II. TFIIE rejoint le complexe pour le stabiliser en se liant à la Pol II, TFIIF, et TBP (Orphanides *et al.*, 1996). Ce complexe va permettre le

recrutement de TFIIF. TFIIE stimule l'activité ADN-hélicase ATP-dépendante et CTD kinase de TFIIF. Ces deux propriétés activent l'ouverture du promoteur et phosphorylent le domaine C terminal de la grande sous-unité de la PolII (CTD) (Ohkuma *et al.*, 1995). TFIIF facilite donc le déroulement de l'ADN simple brin formant la bulle de transcription (Holstege *et al.*, 1996) et la modification de la structure tridimensionnelle de la polymérase qui entraîne la dissociation du complexe d'initiation et le début de la transcription. Enfin TFIIE permet le changement de conformation de la PolII (état ouvert à état fermé) (Westover *et al.*, 2004). Le PIC peut être associé à d'autres complexes multiprotéiques comme le Médiateur qui remplit un rôle d'adaptateur entre les facteurs de transcription et le PIC. Plusieurs types de médiateur sont présents chez l'homme tel que PC2 (Malik *et al.*, 2000) et TRAP (Fondell *et al.*, 1996). Le Médiateur est composé de 26 sous-unités regroupées en 3 domaines (Sato *et al.*, 2004). Le noyau central qui est composé de la tête et du corps médian interagit avec TFIIB-PolII (Dotson *et al.*, 2000). La queue peut interagir avec des activateurs proximaux. Ces trois éléments forment PC2. TRAP est composé de PC2 supplémenté d'un complexe protéique. Les sous-unités interagissent la plupart du temps avec des activateurs comme la sous-unité MED17 avec p53. Le médiateur est connu pour favoriser la formation du PIC soit en augmentant l'efficacité de l'assemblage (Cantin *et al.*, 2003), soit en recrutant directement la PolII ou les facteurs généraux de transcription (Wu *et al.*, 2003).

Facteur général de transcription	Sous-unité	Poids moléculaire	Fonctions
TFIIA	α et β	32 kDa	Impliqués dans la stabilisation de la liaison TBP-ADN; Activation de la transcription
	γ	13 kDa	Activation de la transcription
TFIIB		38 kDa	Fixation à BRE; stabilisation de la liaison TBP-ADN; Implication dans la courbure de l'ADN et la sélection du point +1
TFIIE	a	55 kDa	Interaction et stimulation de TFIIH; Impliqué dans l'ouverture du promoteur
	b	37 kDa	Fixation à l'ADN double brin
TFIIF	RAP 74	82 kDa	Interaction avec l'ADN pendant l'initiation; Positionnement de l'ADN au point +1
	RAP 30	47 kDa	Fixation à l'ARN polymérase II; Implication dans l'élongation
		27 kDa	Commun à TFIID, TFIIF et Swi/Snf
TFIIH	p62 et p52	73 et 59 kDa	Impliqué dans la transcription et à la réparation de l'ADN par excision des nucléotides (NER)
	p34	37 kDa	Impliqué dans la réparation de l'ADN (NER); Rôle structurale dans le complexe TFIIH
	MAT 1	32 kDa	Impliqué dans la réparation de l'ADN (NER); Régulation de l'activité du complexe Cdk7/CycH au sein du complexe CAK (Cdk Activating kinase)
	Cdk 7	35 kDa	Kinase du complexe CAK; Phosphorylation des sérines du CTD de la Pol II
	Cycline H	45 kDa	Cycline de CAK
	XPD/ERCC2	90 kDa	Activité 5'-3' hélicase et ATPase; Impliqué dans la réparation de l'ADN (NER).
	p44	52 kDa	Stimule l'activité hélicase de XPD; Impliqué dans la réparation de l'ADN (NER).
	XPD/ERCC3	95 kDa	Activité 3'-5' hélicase et ATPase dépendante de l'ADN; Ouverture et dégagement du promoteur.

Tableau 1 : Les facteurs généraux de transcription
(adapté de Gangloff *et al.*, 2001)

2. TFIID

La reconnaissance du promoteur par le complexe multi-protéique, TFIID, est primordiale pour l'assemblage du complexe de pré-initiation (PIC) (Müller *et al.*, 2007). TFIID est formé de TBP et de 14 TAFs conservées de la levure à l'homme (*Figure 4*) (Burley et Roeder, 1996). La protéine TBP permet la reconnaissance de la boîte TATA par son

extrémité C-terminale et provoque la courbure de la double hélice d'ADN (Davison *et al.*, 1983).

Les TAFs s'articulent grâce à leur domaine de repliement de type histone (HFD pour Histone Fold Domain) qui augmente les interactions entre les TAFs et stabilise la liaison des TAFs avec l'ADN (Gangloff *et al.*, 2001). Ce domaine comprend trois hélices α reliées par deux boucles. Ce domaine permet l'hétérodimérisation de TAF3/10 (Gangloff *et al.*, 2001), TAF4/12 (Gangloff *et al.*, 2000), TAF6/9 (Bando *et al.*, 1997, Xie *et al.*, 1996), TAF8/10 (Gangloff *et al.*, 2001) et TAF11/13 (Birck *et al.*, 1998, Xie *et al.*, 1996). Ces motifs sont hautement conservés et ont des similitudes avec les histones qui constituent le nucléosome. Les complexes TAF6/9 et TAF4/12 ont pour similitudes H3/H4 et H2A/H2B respectivement (Shao *et al.*, 2005). Ces hétérodimères forment un pseudo-nucléosome (Hoffman et Roeder, 1996). Des études de co-immunoprécipitation et d'immuno-électromicroscopie ont permis de montrer qu'il n'y a qu'une seule copie des TAF1, 2, 7 et TBP et deux copies des TAF5, 10, TAF6/9 et TAF4/12 dans TFIID (Leurent *et al.*, 2002).

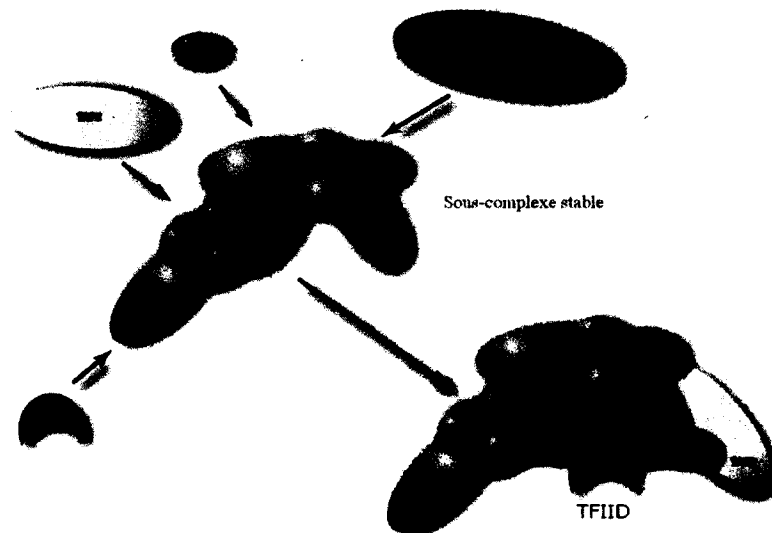


Figure 5 : Le facteur général de transcription TFIID

Le complexe de base composé de TAF4, TAF5, TAF6, TAF9 et TAF12, est complété par TBP, TAF1, TAF2 et TAF11 (tiré de Wright *et al.*, 2006).

Les TAFs sont très importants pour l'activité et l'intégrité de TFIID. En effet les multiples fonctions des TAFs peuvent modifier le comportement de TFIID en fonction des promoteurs existants (*Tableau 2*) (Müller et Tora, 2004). De plus les TAFs peuvent avoir différentes isoformes issues de leurs maturations post-transcriptionnelle ou post-traductionnelle, ce qui est le cas pour TAF4, 6 et 7. Des études ont montré qu'un paralogue de TAF4, TAF4b est plus exprimé dans les lymphocytes B, les testicules et cellules ovariennes (Dikstein *et al.*, 1996, Freiman *et al.*, 2002). Ceci prouve qu'il y a des TAFs spécifiques à certains types cellulaires ou tissus pour l'expression de gènes bien déterminés. Le même constat a été fait pour TAF7 qui est remplacé par TAF7L lors de la spermatogénèse (Pointud *et al.*, 2003). La sous-unité TAF6 possède 5 isoformes après l'épissage alternatif. La forme majoritaire TAF6 α dimérise avec TAF9 et a une activité anti-apoptotique. Par contre TAF6 δ , notre protéine d'intérêt, active l'apoptose et exclut TAF9 de TFIID (Bell *et al.*, 2001). L'absence de 10 acides aminés au centre du domaine de repliement de type histone empêche la dimérisation de TAF6 avec TAF9. TFIID est alors nommé TFIID π car il y a absence de TAF9 (*Figure 5*).

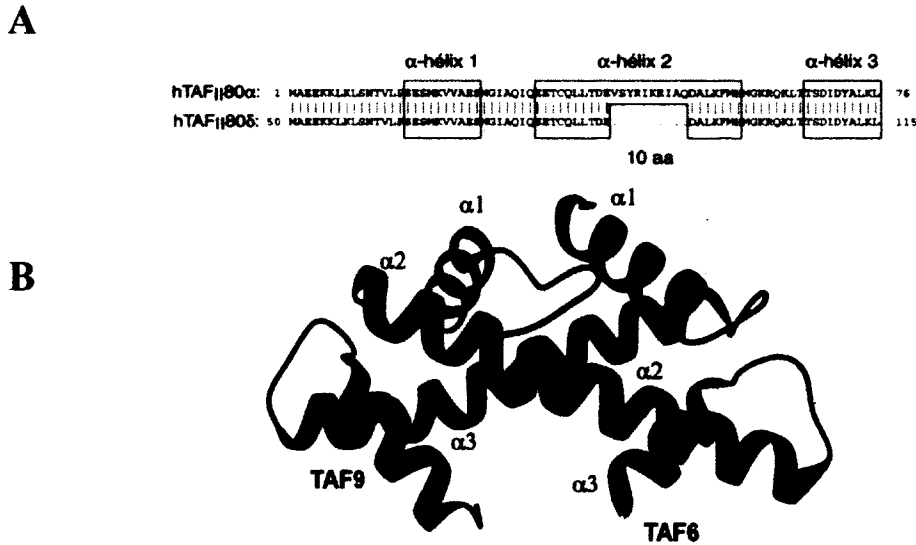


Figure 6 : Le domaine de repliement de type d'histone des variants d'épissage de TAF6 et la dimérisation de TAF6α avec TAF9

A) structure primaire du domaine de repliement de TAF6α et TAF6δ (tiré de Bell *et al.*, 2001) ; B) dimérisation de TAF6α avec TAF9 (tiré de Xie *et al.*, 1996).

Récemment une étude structurale et de biochimie de TAF6 a montré qu'en plus des motifs de type histone, le domaine C terminale de TAF6 interviendrait dans la formation du dimère (TAF6/TAF9) et du trimère (TAF6/TAF9/TAF5) dans TFIID. En effet l'analyse des mutations de TAF6C qui est la deuxième partie du domaine C terminale, démontre que ses 5 domaines de répétitions HEAT participeraient à la dimérisation de TAF6 et TAF9. L'assemblage de TFIID dépendrait aussi de la présence de TAF5, la seule protéine non « histone like », qui consoliderait la formation des lobes de TAFIID (Scheer *et al.*, 2012). Donc la dynamique de TFIID change en fonction de sa composition et de la présence de certains facteurs de transcription. Par conséquent TFIID module l'expression génique dans différents événements physiologiques et adapte la transcription à diverses situations dans les cellules ou tissus.

TAFs	Poids moléculaire	Interactions	Fonctions
TAF1 (TAFII250)	250 kDa	TAF2, Inr	Fixation du promoteur; activités kinase sur TFIIIF, histone acétyltransférase et ubiquitine ligase; impliqué dans le cycle cellulaire
TAF2 (TAFII150)	150 kDa	TAF1, Inr	Fixation du promoteur; impliqué dans le cycle cellulaire
TAF3 (TAFII140)	140 kDa	TAF10	Fixation du promoteur; possède un domaine de repliement d'histone; présence d'une NLS; implication dans la différenciation; inhibition de p53
TAF4 (TAFII130/135)	135 kDa	TAF12	Possède un domaine de repliement d'histone; Interaction avec des co-activateurs (Sp1, c-jun)
TAF5 (TAFII100)	100 kDa	Dimérisation TAF5, TAF6, TAF9	Pilier de TFIIID entre les 3 lobes; impliqué dans le cycle cellulaire
TAF6 (TAFII70/80)	80 kDa	TAF5, TAF9, DPE	Fixation du promoteur; Similaire à l'histone H4
TAF7 (TAFII55)	55 kDa	TAF1	Inhibition des activités acétyltransférase de TAF1 et kinase de TFIIH et P-TEFb
TAF8 (TAFII43)	43 kDa	TAF6, TAF10, TAF11, TBP	Possède un domaine de repliement d'histone; présence d'une NLS; expression pendant la différenciation des adipocytes
TAF9 (TAFII31/32)	32 kDa	TAF5, TAF6	Similaire à l'histone H3; stabilisation de p53
TAF10 (TAFII30)	30 kDa	TAF3, TAF8, TAF13, SPT7Like	Possède un domaine de repliement d'histone; activité cellules spécifiques; inhibition des récepteurs d'œstrogène; impliqué dans la différenciation, le cycle cellulaire
TAF11 (TAFII28)	28 kDa	TAF13, TBP	Similaire à H3; interaction avec TFIIA; stabilisation du complexe TFIIA-TBP-ADN
TAF12 (TAFII15/20)	20 kDa	TAF4, TBP	Similaire à H2B; régulation de l'expression d'E-cadherine
TAF13 (TAFII18)	18 kDa	TAF10, TAF11, TBP	Similaire à l'histone H4
TAF15 (TAFII68)	68 kDa		Sous-unité de TFIIIF, de SWI/SNF, de INO80, de NuA3; interaction avec U1 snRNP et protéines de signalisation cellulaire

Tableau 2 : Les facteurs associés à TBP (TAFs)
(adapté de Gangloff *et al.*, 2001)

C. L'élongation et la terminaison de la transcription

Après une pause de la Pol II, la bulle de transcription entraîne l'échappée de la Pol II du promoteur. Pol II synthétise l'ARN pré-messager qui est complémentaire au brin anti-sens de l'ADN donc cet ARN est de séquence identique au brin sens (brin codant). L'ARN polymérase II a besoin de facteurs d'élongation pour mener à bien la lecture de l'ADN comme TFIIF qui reste lié à la Pol II pour empêcher l'arrêt de l'élongation. Cette étape est dépendante de l'activité kinase de P-TEFb qui phosphoryle la serine 2 du CTD de la Pol II (Marshall *et al.*, 1996). L'hyperphosphorylation du CTD devient une plateforme pour les autres facteurs de transcription et aussi pour des facteurs d'épissage, de polyadénylation. Le facteur ELL et l'élonguine (facteur II) augmentent la vitesse d'élongation de la transcription (Bradsher *et al.*, 1993; Bradsher *et al.*, 1993). Si la PolII venait à s'arrêter, TFIIS (ou facteur II) lié à la grande sous-unité de la Pol II, peut cliver l'ARN et permet la reprise de l'élongation (Nudler *et al.*, 1994).

La Pol II avance le long de l'ADN jusqu'à la transcription du signal de polyadénylation. La reconnaissance de la séquence de polyadénylation par des facteurs de polyadénylation comme CPSF (pour Cleavage and Polyadenylation Specific Factor) ou CStF (Cleavage Stimulation Factor) induit l'arrêt de la transcription, le décrochage de la Pol II et par conséquent la terminaison de la transcription.

Les pré-ARN messagers subissent des modifications co-transcriptionnelles ou post-transcriptionnelles afin que les ARN messagers soit fonctionnels et exportables vers le cytoplasme pour être traduits. L'ARN pré-messager est mûré à trois niveaux. Une coiffe lui est ajoutée en 5' (résidu de la 7 méthyl guanosine). Comme mentionné précédemment le transcrit est polyadénylé à son extrémité 3'. Ces deux étapes de maturation protègent le pré-messager de toute dégradation par les nucléases. Enfin, l'épissage élimine les séquences non codantes entrelacées entre les séquences codantes. Cette étape de maturation sera présentée dans la partie suivante.

III. TAF6

A. TAF6

TAF6 est un facteur de transcription, sous-unité du facteur général de transcription TFIID. La séquence génomique de TAF6, longue de 12 300 paires de base se trouve sur le chromosome 7. Un gène unique mène, par épissage alternatif, à 5 isoformes de TAF6 (*Figure 7*). TAF6 alpha (α) est l'isoforme majoritaire avec 677 acides aminés. TAF6 α ainsi que TAF6 beta (β) et TAF6 gamma (γ) comprennent l'exon deux (exon deux constitutif et alternatif) en entier. TAF6 β (726 acides aminés) possède 49 acides aminés supplémentaires par rapport à TAF6 α car la séquence codante commence par un autre codon ATG qui permet l'insertion des séquences répétitives, les séquences Alus. Par contre TAF6 γ (667 acides aminés) a 10 acides aminés en moins car l'épissage alternatif de TAF6 supprime la partie alternative de l'exon 13. En ce qui concerne TAF6 delta (δ) et epsilon (ϵ), ces deux isoformes ont perdu 10 acides aminés après l'épissage alternatif qui a supprimé la partie alternative de l'exon 2. A la différence de TAF6 δ (667 acides aminés), TAF6 ϵ compte les mêmes 49 acides aminés supplémentaires que TAF6 β . La déplétion de TAF6 chez la drosophile (Aoyagi *et al.* 2001) et Arabidopsis (Lago *et al.* 2005) a montré que cette protéine est essentielle et importante pour le développement des cellules germinales, somatiques et du tube pollinique.

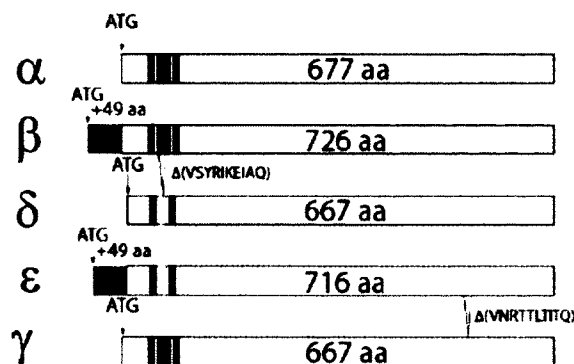


Figure 7 : Les isoformes de TAF6

Les cadres noirs correspondent aux hélices du motif de repliement de type histone ; les cadres gris clair correspondent aux 10 acides aminés manquants après l'épissage alternatif de l'exon 2 ; les cadres gris foncés correspondent aux séquences *Alus* (adapté de Wang *et al.*, 2004).

B. TAF6δ

TAF6δ a été découvert en 2001 par l'équipe du Dr Tora (Bell *et al.*, 2001). Il est issu de l'épissage alternatif de l'exon 2 de l'ARN pré-messager TAF6 par choix de deux sites donneurs d'épissage (5' splice site) (Figure 8). Son activité pro-apoptotique a été mise en évidence après induction de l'apoptose par de l'acide rétinolique (Bell *et al.* 2001). En effet cette caractéristique a été corroborée par le développement d'un oligonucléotide anti-sens dans notre laboratoire (Figure 8). L'oligonucléotide s'hybridant sur l'exon2 alternatif, cache les éléments cis reconnus par les facteurs d'épissage. Ce système force donc l'épissage alternatif en faveur de TAF6δ d'où l'appellation des oligonucléotides anti-sens « splice switching oligonucleotide » (SSO). La transfection de cet oligonucléotide anti-sens dans les cellules HeLa ws a montré que l'induction de TAF6δ active l'entrée des cellules en apoptose (Wilhelm *et al.* 2008). Par ailleurs, nous savons que TAF6δ peut activer l'apoptose indépendamment du suppresseur de tumeur p53, un des facteurs de transcription très étudié pour son activité pro-apoptotique (Wilhelm *et al.* 2010). Des analyses de

micropuces ont été effectuées afin d'observer les profils d'expression du génome après induction de TAF6 δ en comparaison à l'expression de TAF6 α dans des conditions normales de culture cellulaire. Les résultats ont démontré qu'il y avait une nette différence d'expression transcriptomique. En effet, la surexpression de TAF6 δ active significativement les voies de Notch, des intégrines, de l'angiogénèse, de la réponse au stress oxydatif, de p53 en plus de l'apoptose en comparaison à l'expression de TAF6 α (Figure 8). Des données non publiées nous indiquent que les changements d'expression génique par l'induction de TAF6 δ sont bien distincts de ceux intervenant après déplétion de TAF6 par interférence à l'ARN. L'importance de l'impact du changement d'expression de TAF6 α vers TAF6 δ nous amène à nous poser certaines questions, notamment comment l'épissage alternatif en faveur de TAF6 δ est régulé et quelle voie de l'apoptose est activée par l'expression de TAF6 δ .

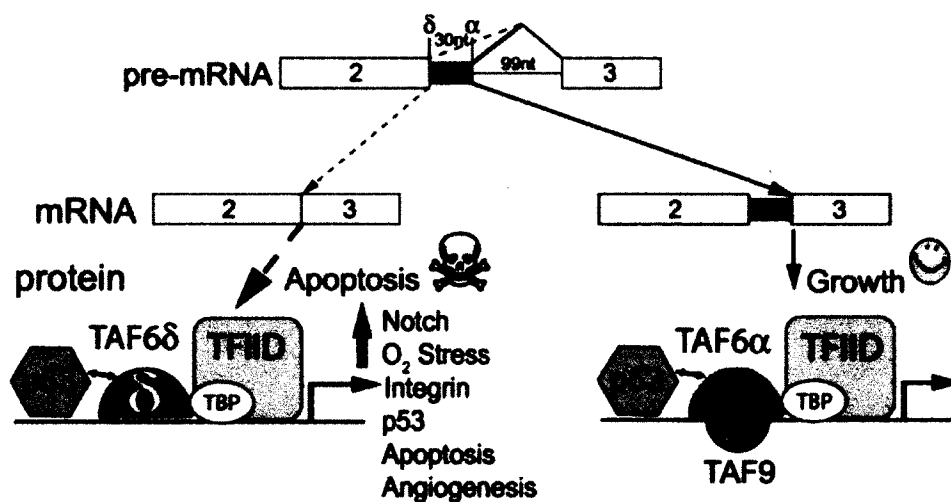


Figure 8 : Le modèle de la voie de TAF6 δ

L'ARN pré-messager subit l'épissage alternatif de son exon 2 qui produit les ARN messagers TAF6 α et TAF6 δ . L'oligonucléotide antisens TAF6 force l'épissage alternatif en faveur de TAF6 δ . La présence de l'exon2 alternatif, le motif central du repliement de type histone favorise la dimérisation de TAF6 α et TAF9. Par contre TAF6 δ ne peut pas interagir avec TAF9 (tiré de Wilhelm *et al.*, 2010).

IV. L'EPISSAGE ET L'EPISSAGE ALTERNATIF DE L'ARN

L'information génétique est constituée de séquences codantes et non-codantes retrouvées dans l'ARN pré-messager nouvellement transcrit. L'épissage est une étape de maturation de l'ARN pré-messager obtenu après la transcription. Les séquences non codantes sont excisées alors que les séquences codantes sont épissées pour obtenir des ARN messenger matures.

L'existence des séquences non codantes appelées introns (INTervening RegiON) a été mis en évidence pour la première fois en 1976-77 lors de l'étude du génome de l'adénovirus de type 2 (Berget *et al.*, 1977, Chow *et al.*, 1977). Les introns sont classés en quatre catégories :

- Les introns autocatalytiques du groupe I (introns des ARN ribosomiques et des gènes des chloroplastes, des mitochondries d'eucaryotes inférieurs)
- Les introns autocatalytiques du groupe II (introns des gènes des mitochondries de levure, de plantes, de chloroplastes)
- Les introns des ARN de transfert
- Les introns des gènes nucléaires (introns des ARN pré-messagers)

A. Les séquences qui définissent les introns

Les introns sont définis par des séquences conservées reconnues par la machinerie d'épissage (*Figure 9*) (Stephens et Schneider, 1992). Il existe deux types d'introns chez les eucaryotes : les introns majeurs, éliminés par le spliceosome U2-dépendant (ou majeur) (Burset *et al.*, 2001) et les introns mineurs éliminés par le spliceosome U12-dépendant (ou mineur), (Jackson, 1991). Dans ce mémoire nous nous focaliserons sur l'épissage des introns par le spliceosome majeur.

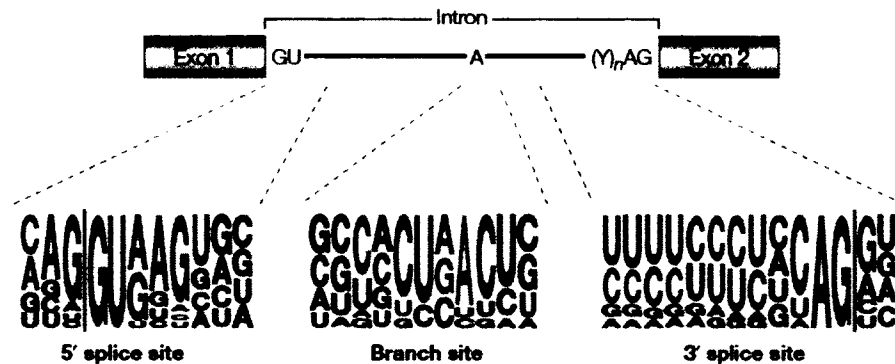


Figure 9 : Les séquences consensus des sites d'épissage

La hauteur des nucléotides représente la fréquence de la présence de ces nucléotides à des positions déterminées. Les nucléotides en bleus sont impliqués dans la reconnaissance de l'intron. L'adénosine en orange est le point de branchement (tiré de Cartegni *et al.*, 2002).

Le site 5' donneur d'épissage (5' Splice Site, 5'SS) est en 5' de l'intron. La comparaison des séquences des introns chez les eucaryotes a permis d'établir une séquence consensus du site 5' donneur d'épissage : AG/GURAGU (R = purine, / = jonction exon/intron). Le dinucléotide GU qui est à la jonction exon-intron, est le plus conservé (Parent et Bisaillon, 2006).

Le site 3' accepteur d'épissage (3' Splice Site, 3'SS) est en 3' de l'intron. Il comprend le site 3'd'épissage proprement dit, le site de branchement et la séquence polypyrimidique. Le site 3' d'épissage caractérisé par une courte séquence YAG/G est à la jonction de l'intron et l'exon suivant. Le site de branchement se trouve 10 à 40 nucléotides en amont du site d'épissage. Sa séquence consensus est plus dégénérée que le site donneur : YNYURAC (Y = pyrimidine, le résidu adénosine correspond au point de branchement). La séquence polypyrimidique est un enchaînement de 10 à 20 résidus pyrimidines (les résidus uraciles) (Parent et Bisaillon, 2006).

B. La réaction chimique de l'épissage

Les introns sont éliminés par deux trans-estérifications successives (*Figure 10*). La première attaque nucléophile vient de l'hydroxyle en 2' du résidu adénosine du site de branchement dans l'intron sur le phosphate en 5' de l'intron. Deux intermédiaires sont formés : le premier exon avec une extrémité 3' hydroxyle libre et un lasso composé de l'intron avec la guanosine 5' terminale lié covalente au site de branchement par une liaison 2'-5' phosphodiester et l'exon 2. La dernière attaque nucléophile de l'hydroxyle en 3' du premier exon sur le phosphate en 3' de l'intron permet la ligation des deux exons et l'élimination de l'intron (Black, 2003).

C. Le spliceosome

L'épissage est réalisé grâce au macrocomplexe multiprotéique, le spliceosome (*Figure 10*). Il est composé de ribonucléoprotéines, les UsnRNP (U1, 2, 4, 5 et 6 snRNP).

Les UsnRNP sont un ensemble de protéines (protéines Sm ou Sm-like et protéines spécifiques) autour d'un ou plusieurs ARN nucléaires riches en uridine (UsnRNA). U1snRNP est composé du snRNA U1 et de 7 protéines Sm, U1-70K, U1A, U1. U2 snRNP est l'assemblage du snRNA U2 avec 7 protéines Sm, U2A', U2B'', les facteurs SF3b, SF3a. Au sein du tri-snRNP U4/U5-U6, U4 et U6 snRNP sont associés. Le duplex d'ARNs est consolidé par les protéines Sm, 15,5K, 61K, 90K. Les protéines 60K, 20K servent de pont entre les 61K et 90K. U5 snRNA est associé aux 7 protéines Sm, U5-15K, 40K, 52K, 100K, 102K, 110K, 116K, 200K et 220K pour former l'U5 snRNP. Ces macro-protéines se lient à des sites bien spécifiques de l'ARN pré-messager au cours de l'épissage.

D. Mécanisme général de l'épissage

L'épissage est une étape de maturation du pré-messager séquentielle ATP et GTP dépendante. L'épissage débute donc par la reconnaissance du site 5' donneur d'épissage par U1snRNP grâce à son snRNA qui reconnaît le pré-messager (jonction exon/intron, -1 ; +5). Les sous-unités U2AF65 et U2AF35 de U2AF s'hybrident respectivement sur la séquence polypyrimidique et sur le site 3' accepteur d'épissage. SF1 reconnaît le site de branchement. La formation du complexe E permet le rapprochement et la définition des exons et introns. En présence d'ATP, le complexe A se forme par des changements d'interactions entre U1, U2 et U2AF. En effet, le pré-spliceosome se forme après la délocalisation de SF1 par U2snRNP sur le site de branchement qui libère le résidu adénosine pour la première attaque nucléophile. Le complexe A devient le complexe B après l'arrivée du tri-snRNP U4 U5 U6. Le spliceosome pré-catalytique libère le site de branchement et le site accepteur. L'interaction de U6 et U2 permet le relargage de U1 et U4. U6 interagit avec la séquence en 5' de l'intron et U5 s'apparie avec le site 3' de l'exon 1. Nouvellement formé, le complexe C (complexe catalytique) active les deux trans-estérification qui permet la ligation des exons et la libération d'un lasso intronique (*Figure 10*) (Krummel *et al.*, 2010).

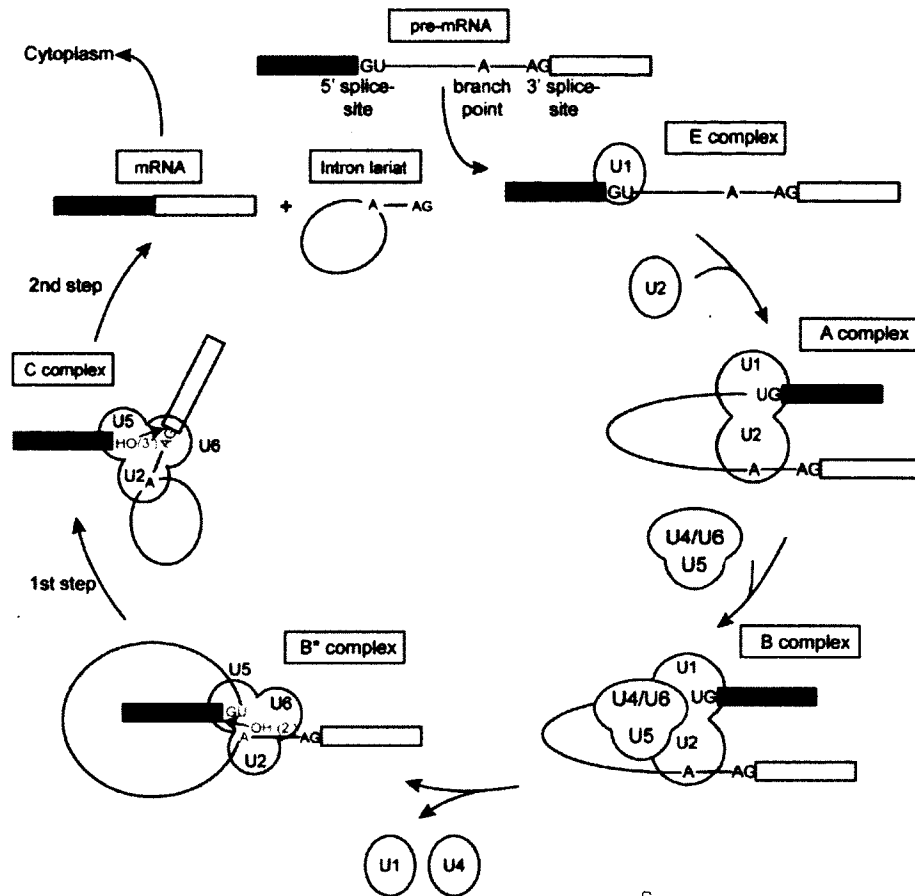


Figure 10 : Le mécanisme de l'épissage

La complexation des UsnRNP sur l'ARN pré-messager et les deux réactions de transestérifications épissent les deux exons et excise l'intron sous forme de lasso (tiré de Krummel *et al.*, 2010).

E. L'épissage alternatif

Grâce à l'épissage alternatif, plusieurs ARN messagers matures proviennent d'un seul ARN pré-messager et par conséquent plusieurs protéines sont traduites à partir d'un seul pré-messager. Cette étape de maturation est à l'origine de la diversité protéique chez les eucaryotes et de l'augmentation de la capacité de codage des gènes. Les messagers résultent de l'utilisation alternative de sites accepteurs et /ou donneurs d'épissage forts ou

faibles pouvant conduire à des protéines différentes. Il existe différents modèles d'épissages alternatifs tels que le saut d'exon, la rétention d'intron, le choix d'exons mutuellement exclusifs, de sites accepteurs et donneurs d'épissage, de promoteurs alternatifs et de sites de polyadénylation alternatifs (Figure 11) (Keren *et al.*, 2010)

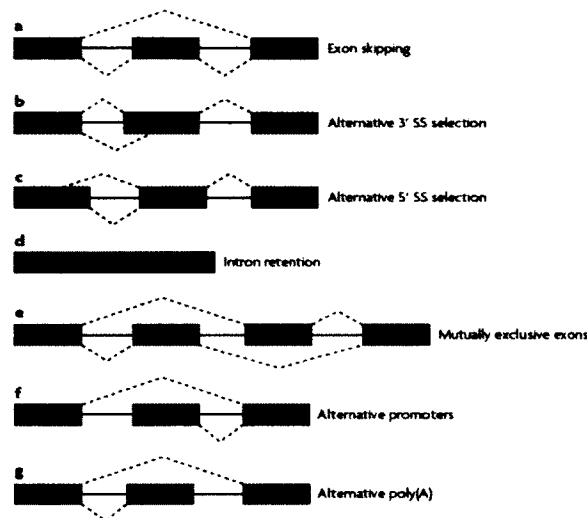


Figure 11 : Les différents types d'épissage alternatif
(tiré de Keren *et al.*, 2010)

De nombreux acteurs interviennent dans la fine régulation de l'épissage alternatif. En effet, la reconnaissance des sites d'épissage n'est pas suffisante pour guider le spliceosome au cours de l'épissage. La présence d'autres séquences *cis* reconnues la plupart du temps par des facteurs d'épissage (facteurs en *trans*) est importante pour le choix des sites donneurs ou accepteurs d'épissage (Figure 12).

1. Les séquences exoniques/introniques et les facteurs d'épissage

Ces séquences se trouvent dans les exons et les introns. Les séquences sont activatrices ou inhibitrices et sont nommées ESE (pour Exonic Splicing Enhancer) ou ISE (pour

Intronic Splicing Enhancer) et ESS (pour Exonic Splicing Silencer) ou ISS (pour Intronic Splicing Silencer) (Cartegni *et al.*, 2002 ; Black, 2003). Elles sont reconnues par deux familles de protéines bien étudiées, les protéines SR et les hnRNPs.

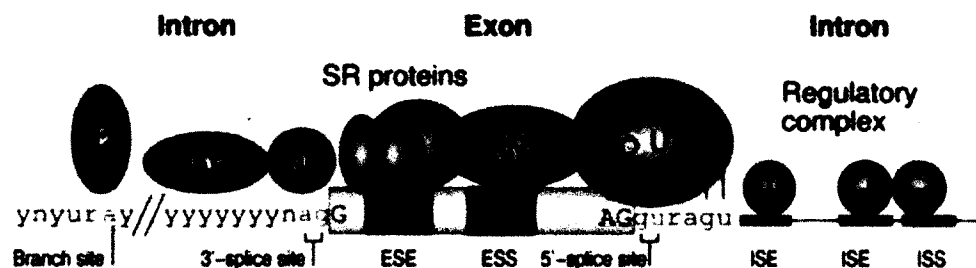


Figure 12 : La régulation de l'épissage alternatif

En plus des séquences consensus des sites d'épissage, des éléments *cis* exoniques et introniques liés par des protéines (protéines SR et hnRNPs) participent au choix des sites d'épissage (tiré de Nilsen et Graveley, 2010).

1.1. Les protéines SR

Les protéines SR sont des facteurs d'épissage qui participent à l'assemblage du spliceosome (Graveley, 2000). Elles possèdent un domaine de liaison à l'ARN (domaine RRM pour RNA Recognition Motifs) et un domaine carboxy-terminal riche en dipeptides arginine et sérine (domaine RS) qui permet les interactions protéine-protéine entre les protéines SR ou avec les protéines du spliceosome. Les protéines SR sont organisées en deux classes. La première comporte les protéines SR contenant le domaine RRM et un domaine RRM associé à un heptapeptide. La deuxième classe n'est composée que d'un seul domaine RRM. Les ESE et ISE sont la plupart du temps liés par les protéines SR (Figure 12).

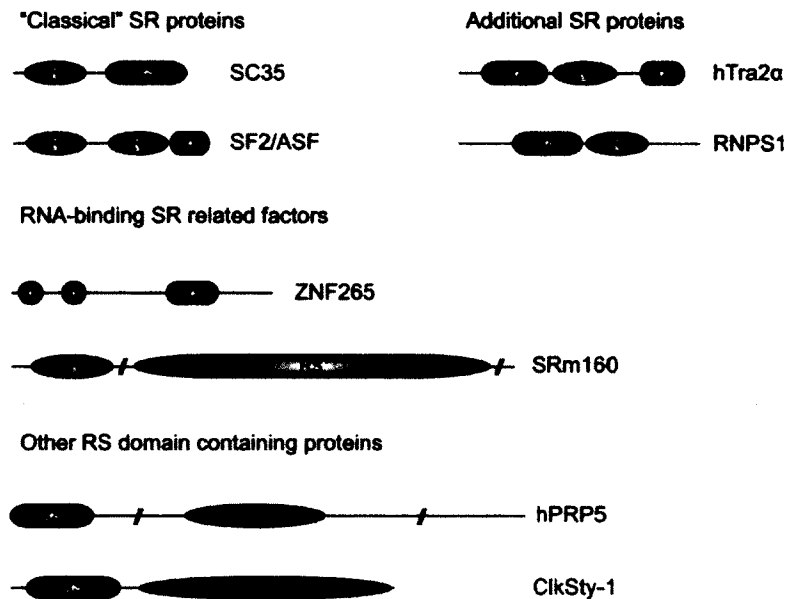


Figure 13 : Les différentes classes de protéines SR

DEAH Box : motif caractéristique des hélicases d'ARN; RS : domaine arginine/serine-rich; PWI: un motif de liaison alternatif d'ARN; Zn : motif de doigts de zinc (tiré de Long et Caceres, 2009).

Des analyses de SELEX ont permis de déterminer les séquences dégénérées reconnues par les protéines SR. Ces facteurs permettent la définition des exons au cours de l'épissage puisque par exemple il a été démontré que ASF et SC35 favorisent la liaison de snRNP U1 au 5'SS, d' U2AF35 au 3'SS et d' U2AF65 à la séquence polypyrimidique. Dans ce cas, les protéines SR consolident la liaison de la machinerie d'épissage à l'ARN pré-messenger. Les protéines SR permettent aussi le rapprochement des deux extrémités des exons lors de la dernière étape de l'épissage ainsi que le recrutement du tri-snRNP. Il existe aussi des protéines SR qui ont une activité inhibitrice. En effet la présence d'ASF/SF2 sur le pré-messenger de SRp20 exclut l'exon 4. Le même phénomène a été observé lorsque SRp30c lié à l'élément CE9, empêche la liaison du site 3' d'épissage et permet l'exclusion de l'exon 8 de hnRNPA1 (Simard *et al.*, 2000 et 2002). Elles participent aussi à l'inhibition de l'activité des hnRNPs qui se lient sur les éléments cis inhibiteurs chevauchant les séquences activatrices.

1.2. Les hnRNPs

Les hnRNP (pour heterogenous nuclear RiboNucleoProtein) sont la deuxième famille des facteurs d'épissage qui se lient aux ESS et aux ISS (*Figure 12*) (Black, 2003). Ces protéines contiennent différents domaines de liaison à l'ARN de type RRM (pour RNA Recognition Motif), RGG (riche en glycine et arginine) ou KH (pour K Homology) et des séquences spécifiques (*Figure 14*). Les hnRNP E, I et K ont une forte affinité pour les séquences polyC, les hnRNP A/B, E, H, H', 2H9, M et U pour les séquences polyG et hnRNP C et PTB pour les séquences polyU. Les protéines de la famille hnRNP A/B s'oligomérisent à un domaine C terminal riche en glycine. Contrairement aux protéines SR, les hnRNP définissent les introns au cours de l'épissage.

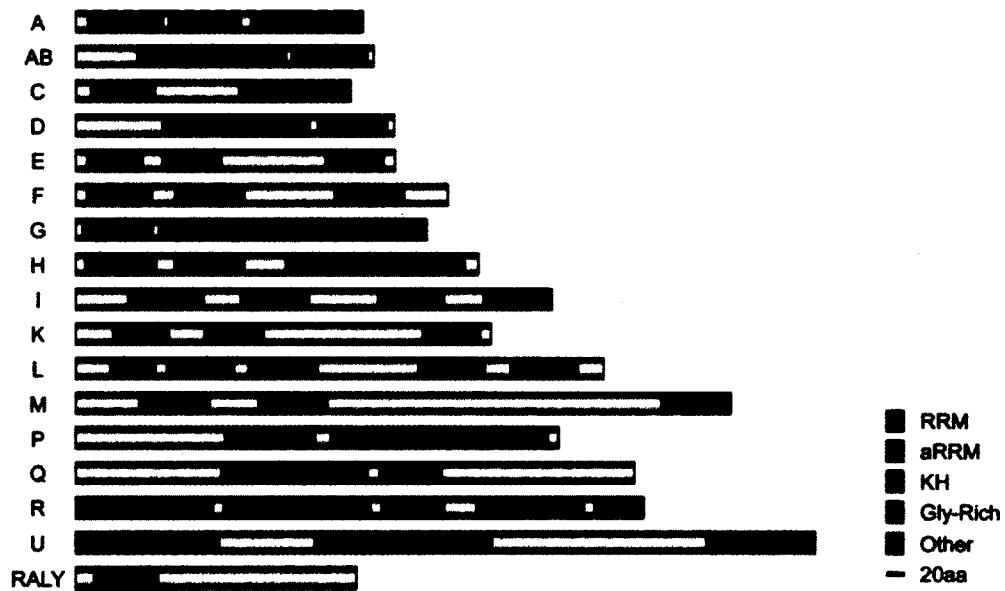


Figure 14 : Les différentes hnRNPs

RRM : motif de reconnaissance de l'ARN ; aRRM : motif de reconnaissance de l'ARN atypique; KH domain: domaine d'homologie K; Gly-Rich domain: domaine glycine riche; other : autre (tiré de Han *et al.*, 2010)

Ils peuvent moduler l'épissage par le rapprochement des introns qui sont très éloignés. Par exemple les hnRNP A1 se lient sur deux introns du pré-messager de hnRNP A1 qui encadrent l'exon 7B. Le rapprochement des deux introns par la liaison des deux hnRNP A1, exclut l'exon 7B de l'ARNm (Chabot *et al.* 2003). Il a été démontré que les hnRNP empêchent la machinerie d'épissage d'accéder aux sites d'épissage. L'encombrement stérique de PTB sur le pré-messager de GABAA α 2 interfère à la reconnaissance du point de branchement et de la séquence polypyrimidique par U2AF65 (Wagner and Garcia-Blanco, 2001). Certains hnRNPs peuvent activer l'épissage de certains pré-messagers comme pour l'épissage de l'exon 6A de la β -tropomyosine de poulet en présence de hnRNP K (Expert-Bezancon *et al.*, 2002), ainsi que pour l'inclusion de l'exon 7 du pré-ARNm smn17 par l'interaction de hnRNP G et hTra2_1 (Hofmann and Wirth, 2002). Il a été mentionné précédemment que les protéines SR avaient pour fonction d'inhiber l'activité des hnRNP. Les hnRNP peuvent aussi inhiber l'activité des protéines SR. Ce phénomène est dû très souvent aux sites des protéines antagonistes qui sont juxtaposés ou chevauchants (House et Lynch, 2008). L'antagonisme est causé par l'encombrement stérique des protéines SR sur les sites de liaison des hnRNP par exemple. Ces interactions sont retrouvées dans les épissages tissus-spécifique (Cáceres *et al.*, 1994 ; Cowper *et al.*, 2001). En effet ces facteurs d'épissage sont exprimés à différents taux selon les tissus concernés. D'autres paramètres peuvent amener un degré supplémentaire à la complexité de la régulation de l'épissage alternatif.

2. La transcription couplée à l'épissage alternatif

L'épissage de l'ARN pré-messager n'est pas un processus isolé mais est une étape co-transcriptionnelle de maturation du pré-messager. Pendant que l'ARN polymérase II synthétise l'ARN pré-messager, la machinerie d'épissage est recrutée grâce à une protéine SR qui interagit en même temps avec la queue CTD de la Pol II (Maniatis et Reed, 2002). Ainsi le pré-messager est épissé en même temps qu'il est synthétisé. Ceci suggère qu'il y aurait une régulation de l'épissage par la transcription. Des études ont

démontré que le ralentissement de la vitesse ou la pause de l'ARN polymérase II donne le temps à l'inclusion d'un exon faible qui est en amont d'un exon fort.

La structure de la chromatine, notamment par les modifications des histones, interviendrait dans la régulation de l'épissage. Des études ont montré que la méthylation des histones dépend des exons (constitutifs, alternatifs, faibles, forts). La marque H3K36me3 se retrouve dans les exons hautement inclus alors que la marque H3K27me2 dans les exons peu inclus (Spies *et al.* 2009). La marque H3K36me3 est plus enrichie dans les exons constitutifs que dans les exons alternatifs (Kolasinska-Zwierz *et al.* 2009). L'inclusion de l'exon alternatif du pré-messager de la β -tropomyosine est fonction de l'augmentation des marques H3K27me3 et H3K4me3 alors que son exclusion dépend de la marque H3K36me3 (Luco *et al.* 2010). Une étude récente a montré que l'épissage est couplé au recrutement de la H3K36 méthyltransférase HYPB au niveau de l'ARN polymérase II lors de l'élongation (De Almeida *et al.* 2011).

3. L'influence des structures secondaires sur l'épissage alternatif

Les ARN pré-messager ont tendance à adopter une structure secondaire locale par des appariements intramoléculaires et interactions hydrophobes. Ces conformations peuvent empêcher l'accès aux sites d'épissage ou à des sites répresseurs d'épissage. La boucle de la tige-boucle contient souvent la séquence d'un site d'interaction aux facteurs d'épissage. L'exemple du pré-messager tau montre que la tige-boucle séquestre le site 5' d'épissage de l'exon 10. Grâce à l'hélicase p68, la structure secondaire est déliée et le site 5' peut être reconnu par U1 snRNP (Kar *et al.*, 2011).

4. La régulation de l'épissage par les ARNs non codants

Récemment l'implication des ARN non codants a été mise en évidence dans la régulation de l'épissage. Le micro-ARN mir-124 qui est exprimé dans les neurones, lie l'ARNm de PTB et active sa dégradation dans les neurones. Mir-124 change l'épissage alternatif en faveur de nPTB par inclusion de l'exon 10. La diminution du taux de PTB dans les neurones active leur différenciation (Makeyev *et al.* 2007). Par contre dans les cellules musculaires, mir-133 empêchant l'expression de nPTB, permet l'épissage en faveur de PTB (exclusion de l'exon 10) qui active la différenciation myoblastique (Boutz *et al.* 2007).

OBJECTIFS

Dans cette introduction, nous avons présenté en lien avec les travaux que nous avons réalisé dans le laboratoire : 1) l'apoptose induite par TAF6 δ , 2) la transcription chez les eucaryotes et l'implication de TAF6 dans l'intégrité et le choix du promoteur au sein de TFIID, 3) l'épissage alternatif chez les vertébrés et sa régulation, enfin 4) notre gène d'intérêt TAF6 et son isoforme TAF6 δ ce que pour laquelle nous menons ce projet.

Au sein de TFIID, TAF6 participe à l'initiation de la transcription. Son absence empêche le bon fonctionnement de l'expression génique pour la formation du tube pollinique chez la plante (Lago *et al.* 2005), au cours de la spermatogénèse ou oogénèse chez la drosophile par exemple (Aoyagi *et al.* 2001). TAF6 est donc essentielle pour la viabilité de ces organismes ainsi que l'homme (Wilhelm *et al.*, 2008). En fonction de l'isoforme de TAF6, TFIID ne se liera pas sur les mêmes promoteurs. En effet, en conditions normales TAF6 α est ubiquitaire et les cellules sont en croissance. Par contre lorsque les cellules expriment TAF6 δ à un seuil supérieur à la normale, elles rentrent en apoptose (Wilhelm *et al.* 2008). TAF6 peut décider de la vie ou de la mort des cellules par l'intermédiaire de ses 2 isoformes. TAF6 active donc une nouvelle voie d'apoptose indépendante de p53 (Wilhelm *et al.*, 2010). La régulation de l'expression de TAF6 est importante pour le devenir des cellules. L'objectif principal du projet est de **caractériser le mécanisme d'épissage alternatif de TAF6, en particulier dans le choix de TAF6 δ** qui constitue une étape majeure pour la décision de vie ou de mort de la cellule.

L'épissage est guidé par la liaison du spliceosome à des séquences qui se conforment de façon variable à des sites consensus définissant les introns des ARN pré-messagers. En plus de ces séquences, des éléments dit cis sont présents dans les exons et introns. Ces éléments sont classés selon leurs propriétés d'activation (ESE, ISE) ou d'inhibition (ESS, ISS) (Ladd et Cooper, 2002). Ils permettent la définition des exons. La première partie de mon projet a été donc de **déterminer les séquences cis** qui régulent l'épissage en faveur de TAF6 δ . Cette thématique avait déjà été étudiée au laboratoire lors de mon arrivée et il

m'a été proposé à la fois de prolonger certaines études déjà entamées par M.-E. Stébenne, mais aussi de compléter la caractérisation des éléments cis.

Les éléments cis sont des séquences courtes (moins de 10 nucléotides) et variables en terme de séquences. Elles sont présentes en plusieurs copies. Ces séquences sont le plus souvent conservées entre les espèces mais avec des motifs dégénérés ce qui les rend difficiles à identifier. Les éléments font office de sites de liaison pour des protéines qui régulent l'épissage en trans. Ces facteurs guident la machinerie d'épissage sur le choix du site 5' d'épissage dans le cas du gène TAF6. La dernière partie de mon projet a été de **développer et mettre au point un nouvel outil**, notamment un vecteur d'expression rapporteur qui nous permettra de déterminer quels sont les facteurs protéiques impliqués dans l'épissage de TAF6.

MATERIELS ET METHODES

1. Lignées cellulaires

Les cellules HeLa ws ont été utilisées et ont été maintenues en culture dans du milieu complet DMEM supplémenté de 2,5% de FBS et de 2,5% de CS (WISENT) avec 5% CO₂ à 37°C.

2. Plasmides

Afin de caractériser les éléments *cis* qui régulent l'épissage de TAF6, les mutations ont été introduites dans le plasmide pcDNA3.1-TAF6. La séquence de TAF6 (249pb) a été amplifiée par PCR à partir du plasmide pGEM-TAF6 (préparé par Dr Bell) avec les oligonucléotides T6-1B (AAAAAGGGATCCCATGGGCATCGCCCAGATTCAGG) et T6-2E (AAAAAGG AATTCCAAGGCGTAGTCAATGTCACTGG) qui contiennent les sites de restrictions BamHI et EcoRI. Le fragment de TAF6 a été inséré dans pcDNA3.1 préalablement digéré par BamHI et EcoRI (New England Biolabs). Ce minigène comprend les exons 2, 3 et l'intron 2 de TAF6.

Le vecteur d'expression pEYFP-TAF6-PTC permettra d'identifier les facteurs d'épissage. Un codon stop a été inséré dans pXJ42-TAF6 (exons 2 à 4 et introns 2 et 3) (Wilhelm *et al.*, 2010) par mutagenèse par PCR avec des oligonucléotides spécifiques (stop 3 +/-, cf. *Tableau 3*). pXJ42-TAF6-PTC a été digéré par XhoI et SmaI (New England Biolabs) pour isoler la séquence TAF6-PTC. Cette séquence a été ensuite insérée dans le plasmide rapporteur pEYFP-N1 qui a été aussi digéré par XhoI et SmaI (New England Biolabs). Le plasmide plentiU6V5A-TAF6-PTC est l'insertion de la séquence TAF6-PTC amplifiée par PCR à partir de pXJ42-TAF6-PTC avec des amorces spécifiques contenant les sites de restriction de XhoI et ApaI. Le vecteur plentiU6V5A-TAF6-PTC sera utilisé pour faire des transfections stables et créer une lignée exprimant EYFP-TAF6-PTC stablement.

Les plasmides pXJ41-hnRNPK (donné gracieusement par Dr. Motorin, Nancy, France) et pGFP-CMV-topaz-hnRNPH (donné gracieusement par l'équipe du Dr. Chabot) ont permis de surexprimer hnRNPK (Marchand *et al.*, 2011) et hnRNPH respectivement.

3. Les oligonucléotides

Les oligonucléotides (IDT) contiennent des mutations flanquées des séquences correspondantes au minigène TAF6. Ces oligonucléotides ont permis d'introduire des mutations dans le minigène TAF6.

Tableau 3 : Liste des séquences des oligonucléotides

PT2.1-RT3	AATAGCGATCCACGCGACTAGTGGGAAGGCACAGTCGAGGCTG
PT2.1	AATAGCGATCCACGCGACTAGTGG
pc3.1Mars	AATACGACTCACTATAGGGAGACCCAAGC
T6-1B	AAA AAG GGA TCC CAT GGG CAT CGC CCA GAT TCA GG
T6-2E	AAA AAG GAA TTC CAA GGC GTA GTC AAT GTC ACT GG
Pcons + (alpha)	AA GAG ATC GCA CAG GTA AGT CGG CCC TTC TGT CC
Pcons -	GG ACA GAA GGG CCG ACT TAC CTG TGC GAT CTC TT
Dcons + (delta)	A ACG GAT GAG GTA AGT TAC CGC ATC AAA GAG ATC GC
Dcons -	GC GAT CTC TTT GAT GCG GTA ACT TAC CTC ATC CGT T
Branch pt ko+	G ATT TTG GGC TGG TGT CTG TTC TGT CAC C
Branch pt ko-	G GTG ACA GAA CAG ACA CCA GCC CAA AAT C
ex2 1+	ATG GGC ATC ATT TGA GCC TAG GAG GAG AC
ex2 1-	GT CTC CTC CTA GGC TCA AAT GAT GCC CAT
ex2 2+	GCC CAG ATT CGT TGT TGT GAC TGC CAG C
ex2 2-	G CTG GCA GTC ACA ACA ACG AAT CTG GGC
ex2 3+	AG GAG GAG ACA CTA ACT AGA CTA ACG GAT GAG G
ex2 3-	C CTC ATG CGT TAG TCT AGT TAG TGT CTC CTC CT
ex3-1 +	CCC CAG GAT CGG AAC TTC AAC ATG CAC ATG GG
ex3-1 -	CC CAT GTG CAT GTT GAA GTT CCG ATC CTG GGG
ex3-2 +	GCC TTG AAG TTG TAC GTG TAG GGG AAG CGG
ex3-2 -	CCG CTT CCC CTA CAC GTA CAA CTT CAA GGC
ex3-3 +	ATG CAC ATC CCC TTC GCC GAG AAG CTC ACC
ex3-3 -	GGT GAG CTT CTC GGC GAA GGG GAT GTG CAT
C₃-91+	CA CAG GTG ACC CGG AAA TTC TGT CCA GC

C ₃ -91-	GC TGG ACA GAA <u>TTT</u> CCG GGT CAC CTG TG
C ₃ -125+	CT CTA CAT TGG GTC CAT <u>AGT GAA</u> CAG GGC CCC G
C ₃ -125-	C GGG GCC CTG <u>TTC ACT</u> ATG GAC CCA ATG TAG AG
C ₄ -132 +	TC CAT GTC ACC CAG GGA <u>AAA</u> GTC CTT AGG
C ₄ -132 -	CCT AAG GAC <u>TTT TCC</u> CTG GGT GAC ATG GA
C ₃ -125 ; C ₄ -132 +	AT TGG GTC CAT GTC <u>AAA AAG</u> GGA <u>AAA</u> GTC CTT AGG
C ₃ -125 ; C ₄ -132 -	CCT AAG GAC <u>TTT TCC</u> CTT <u>TTT</u> GAC ATG GAC CCA AT
C ₃ -175 +	G TTC TGT CAC CGA <u>TTT</u> CAG GAT GCC
C ₃ -175 -	GGC ATC CTG <u>AAA</u> TCG GTG ACA GAA C
G ₃ -113 +	CCA GCT CTA CAT <u>TCA TTC</u> CAT GTC ACC C
G ₃ -113 -	G GGT GAC ATG GAA TGA ATG TAG AGC TGG
G ₃ -129 +	CAT GTC ACC CAC <u>ATC</u> CCC GTC CTT AGG
G ₃ -129 -	CCT AAG GAC GGG <u>GAT GTG</u> GGT GAC ATG
G ₃ -150 +	C CCC GTC CTT AGG ATT <u>TTC ATC</u> TGA TGT CTG
G ₃ -150 -	CAG ACA TCA <u>GAT</u> GAA AAT CCT AAG GAC GGG G
A68C +	C TAC CGC ATC <u>ACA</u> GAG ATC GCA CAG G
A68C -	C CTG TGC GAT CTC <u>TGT</u> GAT GCG GTA G
A73G +	AAA GAG <u>GTC</u> GCA CAG GTG ACC CGG
A73G -	CCG GGT CAC CTG TGC <u>GAC</u> CTC TTT
A73C +	AAA GAG <u>CTC</u> GCA CAG GTG ACC CGG
A73C -	CCG GGT CAC CTG TGC <u>GAG</u> CTC TTT
A73U +	AAA GAG <u>ITC</u> GCA CAG GTG ACC CGG
A73U-	CCG GGT CAC CTG TGC <u>GAA</u> CTC TTT
C86U +	GCA CAG GTG <u>ATC</u> CGG CCC TTC TGT CC
C86U -	GG ACA GAA GGG CCG <u>GAT</u> CAC CTG TGC
WT74-77+	ATC AAA GAG AAA <u>CGA</u> CAG GTG ACC CGG
WT74-77-	CCG GGT CAC CTG <u>TCG TTT</u> CTC TTT GAT
WT74-75+	ATC AAA GAG AAA <u>GCA</u> CAG GTG ACC CGG
WT74-75-	CCG GGT CAC CTG TGC <u>TTT</u> CTC TTT GAT
WT76-77+	ATC AAA GAG ATC <u>CAA</u> CAG GTG ACC CGG
WT76-77-	CCG GGT CAC CTG <u>TTG</u> GAT CTC TTT GAT
WT75+	ATC AAA GAG ATT <u>GCA</u> CAG GTG ACC CGG
WT75-	CCG GGT CAC CTG TGC <u>AAT</u> CTC TTT GAT
WT78+	ATC AAA GAG ATC <u>GCT</u> CAG GTG ACC CGG
WT78-	CCG GGT CAC CTG <u>AGC</u> GAT CTC TTT GAT
GC74-77+	ATC AAA GAG <u>GAA</u> CGA CAG GTG ACC CGG
GC74-77-	CCG GGT CAC CTG <u>TCG TTC</u> CTC TTT GAT
GC76-77+	ATC AAA GAG <u>GTC</u> <u>CAA</u> CAG GTG ACC CG
GC76-77-	CGG GTC ACC TGT <u>TGG</u> <u>ACC</u> TCT TTG AT
GC75+	ATC AAA GAG <u>GTT</u> GCA CAG GTG ACC CG

GC75-	CCG GGT CAC CTG TGC <u>AAC</u> CTC TTT GAT
GC78+	ATC AAA GAG <u>GTC</u> GCT <u>CAG</u> GTG ACC C
GC78-	GGG TCA CCT <u>GAG</u> CGA <u>CCT</u> CTT TGA T
ACA +	C TGC CAG CTG CTA ACG <u>ACA</u> GAG GTC AGC TAC CGC
ACA -	GCG GTA GCT GAC CTC <u>TGT</u> CGT TAG CAG CTG GCA G
CUC +	C TGC CAG CTG CTA ACG <u>CTC</u> GAG GTC AGC TAC CGC
CUC -	GCG GTA GCT GAC CTC <u>GAG</u> CGT TAG CAG CTG GCA G
UUG +	G CTG CTA ACG <u>CAA</u> GAG GTC AGC TAC CGC <u>TTG</u> AAA GAG ATC GC
UUG -	GC GAT CTC TTT <u>CAA</u> GCG GTA GCT GAC CTC <u>TTG</u> CGT TAG CAG C
A64U +	GAG GTC AGC TAC CGC <u>ITC</u> AAA GAG ATC GCA C
A64U -	G TGC GAT CTC TTT <u>GAA</u> GCG GTA GCT GAC CTC
C66U +	G GTC AGC TAC CGC ATT <u>AAA</u> GAG ATC GCA CAG G
C66U -	C CTG TGC GAT CTC TTT <u>AAT</u> GCG GTA GCT GAC C
C66A +	A ACG GAT GAG GTC AGC TAC CGC <u>ATA</u> AAA GAG ATC GC
C66A -	GC GAT CTC TTT <u>TAT</u> GCG GTA GCT GAC CTC ATC CGT T
C66G +	A ACG GAT GAG GTC AGC TAC CGC <u>ATG</u> AAA GAG ATC GC
C66G -	GC GAT CTC TTT <u>CAT</u> GCG GTA GCT GAC CTC ATC CGT T
86C-A +	CGCACAGGTGA <u>ACCGGCCCTTCTGTCC</u>
86C-A -	GGACAGAAGGGCCGG <u>ITCACCTGTGCG</u>
TAF6-EYFP reverse PCR	GG TGG ATC CCG GGG CAG AGG GGT ATT
EYFP reverse RT	TTT ACG TCG CCG TCC AGC TCG ACC
pE-TAF6S-XL	<u>TCGAG</u> ATG GCT GAG GAG AAG AAG CTG
pE-TAF6S-XS	<u>G</u> ATG GCT GAG GAG AAG AAG CTG AAG C
pE-TAF6S-ApL	<u>GGCCC</u> TTA CTT GTA CAG CTC GTC CAT GCC G
pE-TAF6S-ApS	<u>C</u> TTA CTT GTA CAG CTC GTC CAT GCC G
stop1 +	AGC TAC CGC ATC <u>TAA</u> GAG ATC GCA CAG G
stop1 -	C CTG TGC GAT CTC <u>TTA</u> GAT GCG GTA GCT
stop2 +	AG GTC AGC TAC CGC ATC AAA <u>TAG</u> ATC GCA CAG G
stop2 -	C CTG TGC GAT CTA <u>TTT</u> GAT GCG GTA GCT GAC CT
stop3 +	AT GAG GTC AGC TAG <u>CGC</u> ATC AAA GAG ATC GC
stop3 -	GC GAT CTC TTT GAT GCG <u>CTA</u> GCT GAC CTC AT
stop4 +	A ACG GAT GAG GTC AGC TAC <u>TGA</u> ATC AAA GAG ATC GC
stop4 -	GC GAT CTC TTT GAT <u>TCA</u> GTA GCT GAC CTC ATC CGT T

4. Les anticorps

Nous avons utilisé l'anticorps monoclonal 37TA1C2 qui est dirigé contre TAF6 δ (Bell *et al.*, 2001). Afin de confirmer la surexpression de hnRNPK dans les cellules HeLa un anticorps anti-hnRNPK de l'équipe de Dr. Chabot a été utilisé. Un anticorps anti-GFP (Roche) a été utilisé pour observer l'expression d'EYFP-TAF6-PTC qui a le même épitope. Cet anticorps a été aussi utilisé contre hnRNPH dont la séquence est fusionnée à GFP (pGFP-CMV-topaz-hnRNPH). L'anticorps anti- α -tubuline (clone B4-4-1-2, Sigma) a permis de contrôler les quantités de protéines dans chaque piste des immunobuvardages de type western. Un anticorps secondaire anti-souris (Goat anti-mouse HRP, Jackson Immuno Research Laboratories inc.) a été choisi pour la révélation des immunobuvardages de type Western.

5. Mutagenèse par réaction en chaîne par polymérase (PCR)

Les mutations désirées ont été introduites dans le minigène TAF6 par PCR grâce aux paires d'oligonucléotides (IDT) (cf. *Tableau 3*).

Cinquante nanogrammes du plasmide pcDNA3.1-TAF6 ont été amplifiés par PCR avec 1X du tampon de l'ADN polymérase Pfu, 200 μ M de chaque dNTPs (Fermentas), 200 nM de chaque oligonucléotides et de 0.5 U de l'ADN polymérase Pfu dans un volume final de 50 μ L. La réaction de PCR est composée de la phase de dénaturation initiale à 95°C pendant 3 minutes, de dénaturation à 95°C pendant 30 secondes, de la phase d'hybridation à une température d'hybridation qui dépend des oligonucléotides pendant une minute, de la phase d'élongation à 72°C pendant 6 minutes. Ces trois phases ont été répétées 18 à 20 fois selon les mutations. La PCR se termine par une phase d'élongation finale à 72°C pendant 5 minutes. Les produits de PCR ont été par la suite digérés par l'enzyme de restriction DpnI (New England Biolabs) pour éliminer le plasmide matrice qui est méthylé et donc sensible à cette enzyme contrairement aux séquences amplifiées.

par PCR. Les plasmides nouvellement synthétisés ont transformés les bactéries chimio-compétentes *E. coli* XL1 bleues ou DH5 α .

6. Transformation bactérienne

Pour avoir une plus grande quantité les plasmides mutés ont été amplifiés par culture bactérienne.

Un mélange réactionnel comprenant 10 μ L des plasmides nouvellement synthétisés, dans un volume final de 100 μ l de KCM 1X (KCl 0.1 M, CaCl₂ 0.03 M et MgCl₂ 0.05 M) a été incubé avec 100 μ L de bactéries chimio-compétentes (*E. coli* XL1 bleues ou DH5 α) 20 minutes sur glace et 10 minutes à température ambiante. Afin qu'il y ait croissance des bactéries transformées par le produit de PCR, il a été ajouté 1 mL de milieu de culture LB (Wisent). Après une incubation d'une heure sous agitation à 37°C, le mélange de transformation a été centrifugé à 3000 rpm pendant 3 minutes. Le surnageant est retiré pour ne laisser qu'environ 100 μ l dans lesquels le culot de bactéries est resuspendu et étalé sur une boîte de pétri LB agar (Bioshop) contenant l'antibiotique, l'ampicilline (100 μ g/mL) (Wisent). Les bactéries étalées ont été incubées à 37°C la nuit.

7. Mini-préparation des ADN plasmidiques

Cette étape d'isolement et de purification des plasmides a été faite à l'aide des kits « EZ-10 spin column plasmid DNA » de la compagnie BIOBASIC INC. La veille de la mini-préparation, plusieurs colonies ont été piquées et inoculées dans 2 mL de milieu LB contenant de l'ampicilline (100 μ g/mL) et incubées toute la nuit à 37°C. Les trois quarts de la culture ont été centrifugés 2 minutes à 12 000 rpm. Le culot de bactéries a été resuspendu dans 100 μ L du premier tampon qui contient de la RNase A pour détruire les ARNs contaminants. Il a été ajouté 200 μ L du tampon II qui a permis de lyser les

bactéries pour libérer l'ADN plasmidique. Trois cents microlitres du tampon III ont neutralisé le tampon II. Le tout a été déposé sur une colonne puis centrifugé à 12 000 rpm pendant 5 minutes. La colonne a été lavée et centrifugée à 10 000 rpm, 2 minutes par un tampon de lavage à deux reprises. Une dernière centrifugation a éliminé les éventuelles traces d'éthanol sur la colonne. L'ADN plasmidique a été élué par 50 μ L de tampon d'élution après centrifugation de deux minutes. Le plasmide a été digéré par des enzymes de restriction spécifiques à la caractérisation de ce plasmide. Ensuite l'ADN plasmidique a été conservé à -20°C en attendant les résultats des séquençages de la plateforme de séquençage et de génotypage de l'Université de Laval. Ces résultats nous ont permis de vérifier la présence des mutations attendues et éventuellement des mutations accidentelles apparues au cours de la mutagenèse par PCR.

8. Midi-préparation des ADN plasmidiques

Une pré-culture de 2 mL de LB-ampicilline (100 μ g/mL) et de 200 μ L de la mini-préparation d'intérêt a été incubée toute la journée. Par la suite, 1 mL de la pré-culture a été mélangé avec 100 mL de milieu LB-ampicilline (100 μ g/mL) la nuit. Le kit « HiPure plasmid midiprep » de INVITROGEN (catalog no. K2100-04) nous a permis de faire la midi-préparation des ADN plasmidiques. Toute la midi-préparation a été centrifugée à 4 000 g pendant 10 minutes. Le culot a été resuspendu par 6 mL de la solution R3 qui contient de la RNase A. Six millilitres de la solution L7 a permis de lyser les bactéries. Après incubation à température pièce 5 minutes, 6 mL de la solution N3 ont été ajoutés pour précipiter l'ADN génomique tout en neutralisant la solution L7. Le mélange a été centrifugé 10 minutes à 12 000 g afin d'isoler le plasmide dans le surnageant. Ce surnageant a été déposé sur une colonne échangeuse d'ions après son équilibration par 10 mL de la solution EQ1. La colonne a été lavée deux fois par 10 mL de la solution W8 pour décrocher les produits qui sont élués à faible concentration d'ions. L'ADN plasmidique est élué de la colonne par 5 mL de la solution d'élution. L'éluat a été

précipité par 3.5 mL d'isopropanol puis centrifugé 45 minutes à 15 000 g. Le culot obtenu a été lavé par 3 mL d'éthanol 70%. Le culot a été séché à l'air libre pour enlever toutes traces d'éthanol avant d'avoir été resuspendu dans 200 µL de la solution TE (10 mM Tris-Cl, pH 8.0 – 1 mM EDTA). La midi-préparation a aussi subi une digestion analytique avec des enzymes spécifiques et elle a été séquencée également. L'ADN plasmidique est conservé à -20°C. Ces plasmides contenant les mutations ont été utilisés pour effectuer les transfections des cellules HeLa ws.

9. Transfection transitoire des cellules HeLa ws

La veille d'une transfection dans une plaque 24 puits, 70 000 cellules HeLa ws ont été réparties par puits dans 500 µL de milieu complet. Elles ont été mises en culture 24h à 37°C et 5% CO₂ pour qu'elles deviennent confluentes à temps pour la transfection. Le jour de la transfection, le milieu complet a été remplacé par 300 µL d'opti-MEM (Wisent) et mis à incuber à 37°C et 5% CO₂, le temps de la préparation des mélanges de transfection. Un microlitre de DMRIE-C (le transfectant) (Invitrogen) a été mélangé à 150 µL d'opti-MEM. Cette solution a été incubée 30 minutes à température ambiante. Pendant cette incubation, 250 ng d'ADN plasmidique ont aussi été mélangés avec 150 µL d'opti-MEM. Par la suite les 150 µL d'opti-MEM/DMRIE-C ont été incorporés au 150 µL d'opti-MEM/ADN plasmidique. Le tout a été incubé 15 minutes à température pièce pour favoriser la formation des complexes ADN/DMRIE-C. Ce mélange a été déposé dans les puits de la plaque à la place des 300 µL d'opti-MEM qui ont été mis au début de l'expérience. Les cellules en présence du complexe ADN/DMRIE-C ont été incubées à 37°C, 5% CO₂ pendant 24 heures. L'opti-MEM qui a servi à la transfection a été remplacé par le milieu complet. Par la suite l'incubation des cellules a été prolongée de 18 heures dans les conditions standards.

10. Transfection des cellules HeLa ws par les oligonucléotides antisens : les SSOs (splice switching oligonucleotides) et les siRNA (silencing RNA)

Le milieu a tout d'abord été remplacé par 350 μ L de milieu complet. La lipofectamine (Invitrogen) qui permet la transfection des cellules HeLa ws par les petits ARNs, pour un volume de 0,8 μ L a été mélangé avec 50 μ L d'opti-MEM par puits et ce mélange a été incubé 5 à 10 minutes à température pièce. Par la suite, les petits ARN ont été mélangés à 50 μ L d'opti-MEM pour avoir une concentration finale de 100 nM par puits transfectés. La suspension de lipofectamine a été combinée à celle des ARN délicatement. Les complexes se sont formés pendant 20 minutes minimum à température pièce. Les 100 μ L de complexe ont été ajoutés à la surface des 350 μ L du milieu complet pour chaque puits. Enfin, la plaque 24 puits a été incubée pendant 18 heures supplémentaires.

11. Extraction d'ARN

Les cellules transfectées ont été récoltées. Le milieu DMEM a été retiré et transféré dans un eppendorf de 1.5 ml. 250 μ L de trizol (Invitrogen) ont été ajoutés dans les puits pour pouvoir récupérer les cellules transfectées adhérentes aux puits. La plaque a été incubée 2-3 minutes à 37°C. Pendant ce temps le milieu a été centrifugé 30 secondes à 10 000 rpm. Le trizol a été ajouté au culot de cellules obtenu après centrifugation du milieu. 50 μ L de chloroforme ont été ajoutés au lysat de trizol. La solution a été mélangée énergiquement et incubée 2 minutes à température ambiante. Une centrifugation de 15 minutes à 13 000 rpm a permis de séparer la phase aqueuse contenant les ARNs de la phase organique. La phase aqueuse a été prélevée et mélangée par inversement à 125 μ L d'isopropanol qui précipite les ARNs après une centrifugation de 15 minutes à 13 000 rpm. Le culot d'ARN obtenu a été lavé avec 1 mL d'éthanol 70%. Après une centrifugation de 5 minutes à 8 000 rpm, le culot a été séché à l'air libre. Le culot bien

sec a été solubilisé dans 3 à 15 μ L d'eau nano stérile selon sa grosseur. Les ARNs sont conservés à -80°C.

12. Transcription inverse (RT)

Un microgramme d'ARN a servi pour la transcription inverse. Les ARN extraits provenant des transfections des plasmides pcDNA3.1-TAF6 sauvage ou muté et de pEYFP-TAF6-PTC ont été traités à la DNase (1 μ L) (Promega) complété avec 1 μ L de tampon DNase et d'eau nano stérile pour un volume final de 10 μ L pendant 30 minutes à 37°C. L'enzyme digérant d'éventuels ADN contaminants a été inactivée par l'ajout de 1 μ L de tampon stop pendant 10 minutes à 65°C. Par contre les transfections des plasmides pGFP-CMV-topaz-hnRNPH et pXJ41-hnRNPK n'ont pas subi cette étape de digestion. 3.3 μ M d'oligodT ou des oligonucléotides PT2.1-RT3 (AATAGCGATCCACGCGACT AGTGGGAAGGCACAGTCGAGGCTG), EYFP reverse RT (TTTACGTCGCCGTCC AGCTCGACC) ont été mélangés aux ARNs pendant 2 minutes à 95°C pour qu'ils s'hybrident en 3' des ARNs néo-synthétisés. Les oligodT, les oligonucléotides PT2.1-RT3 et EYFP reverse RT s'hybrident spécifiquement à TAF6 endogène, TAF6 exogène provenant de la transfection de pcDNA3.1-TAF6 et pEYFP-TAF6-PTC respectivement.

La transcription inverse se fait donc avec un microgramme d'ARN dans 12 μ L auxquels on ajoute la solution de transcription contenant 1X de tampon MMuLV, 1mM de dNTPs, 5 μ M de dithiothreitol (DTT), et 0.0015 U de la transcriptase inverse MMuLV (produite par Catherine Desrosiers) et de l'eau pour avoir un volume final de 20 μ L. La transcription inverse s'est faite pendant 1 heure à 42°C. Les ADNc ont été conservés à -80°C.

13. Réaction en chaîne par polymérase (PCR)

L'ADN complémentaire (ADNc) obtenu après transcription inverse d'ARN sert de matrice pour une amplification par réaction en chaîne par polymérase. Deux microlitres d'ADNc ont été mélangés dans une solution de 10 µl avec 1X du tampon de l'ADN polymerase Taq, 200 µM de chaque dNTPs, 400 nM d'oligonucléotides (IDT) sens : T6-1B (AAAAAGGGATCCCATGGGCATCGCCCAGATTCAGG) et antisens : T6-2E (AAAAAGGAATTCCAAGGCGTAGTCAATGTCACTGG) (ADNc TAF6 endogène), PT2.1 (AATAGCGATCCACGCGACTAGTGG) (ADNc TAF6 exogène pcDNA3.1-TAF6), TAF6-EYFP reverse PCR (GGTGGATCCCGGGGCAGAG G GGTATT) (ADNc TAF6 exogène pEYFP-TAF6-PTC), 1.25 U de l'ADN polymerase Taq (produite par Catherine Desrosiers). La réaction de PCR est composée de la phase de dénaturation initiale à 94°C pendant 3 minutes, de dénaturation à 94°C pendant 1 minute, de la phase d'hybridation à 58°C pendant 45 secondes, de la phase d'élongation à 68°C pendant 50 secondes. Ces trois phases ont été répétées 25 fois. La PCR se termine par une phase d'élongation finale à 68°C pendant 5 minutes. Les produits de PCR sont conservés à -20°C.

14. Quantification qualitative des produits de PCR

Les produits de PCR ont été analysés qualitativement par gel d'électrophorèse de polyacrylamide 15% non dénaturant afin d'observer les patrons d'épissage alternatif de TAF6. Les gels de polyacrylamide sont un mélange de 15% d'acrylamide : bis acrylamide (30 : 0.8), de TBE 1X (Tris base 0.09 M, EDTA 0.1 M et acide borique 0.089 M), 42 µl d'ammonium persulfate (APS), 2,2 µl de TEMED et d'eau pour avoir une solution de 6 ml. Cinq microlitres de produits de PCR ont migré dans un gel d'électrophorèse de polyacrylamide. Le gel a été coloré au bromure d'éthidium (BET) dilué par du tampon TBE 1X. Ainsi une photographie a pu être prise aux UV.

15. Quantification des produits de PCR

Les produits de PCR ont été également analysés quantitativement par électrophorèse en capillarité, au moyen du Bioanalyser 2100 (Agilent). Les ADNc amplifiés ont été déposés dans des puces à ADN. Les produits de PCR migrent en fonction de leur taille et charge. Les résultats de la migration nous sont donnés sous forme d'une courbe avec différents pics qui représentent les divers isoformes d'épissage. Le pourcentage des isoformes a été calculé en fonction de la quantité totale de variant d'épissage ($[\delta/(\alpha+\delta)]$).

16. Quantification de l'expression de TAF6 par la fluorescence des cellules transfectées par pEYFP-TAF6-PTC par cytométrie de flux

Les cellules HeLa ws ont été transfectées par 1) 25 ng de pEYFP-TAF6-PTC, 2) pEYFP-TAF6-PTC et 10 nM de l'oligonucléotide antisens ExonB (SSO ExonB), 3) pEYFP-TAF6-PTC et l'oligonucléotide antisens contrôle (SSO Ctl), 4) 25 ng de pEYFP-N1 et 5) 25 ng de pEYFP-TAF6. Les différentes transfections ont été complétées par 75 ng de pBSK pour avoir un minimum de 100 ng d'ADN transfectant. Pour observer l'émission de fluorescence produite par l'expression d'EYFP-TAF6-PTC, les cellules transfectées ont été récoltées et analysées en cytométrie de flux.

Le milieu complet des puits a été prélevé. Les cellules ont été lavées avec 450 μ L de PBS 1X qui est à température ambiante. Le PBS 1X a été prélevé et mélangé avec le milieu complet. Les aliquots ont été conservés sur glace. Pour récolter les cellules adhérentes, 45 μ L de trypsine ont été déposés dans les puits. La plaque a été mise à incuber quelques minutes à 37°C. Pour décoller les cellules des puits, 300 μ L de milieu complet ont été ajoutés dans les puits et des aller - retour de pipetage ont permis de prélever les cellules. Le milieu contenant les cellules a été transféré dans les tubes utilisés précédemment sur glace. Les puits ont été rincés au PBS 1X tempéré. Les tubes ont été centrifugés deux fois 2 minutes à 4 000 rpm à 4°C. Le surnageant a été enlevé pour laver le culot à nouveau.

avec 500 μL de PBS 1X et le centrifuger deux fois 2 minutes à 4 000 rpm à 4°C. Enfin le culot a été resuspendu dans 100 à 200 μL de PBS 1X. les échantillons ont été analysés par cytométrie de flux pour observer la quantité de fluorescence des cellules, donc l'expression de TAF6 δ .

17. Immunobuvardage de type Western

Les cellules transfectées ont été récoltées sur glace afin de préserver les protéines. Elles ont été lavées par 500 μL de PBS 1X (NaCl 137 mM, KCl 2.7 mM, Na_2HPO_4 4.3 mM et KH_2PO_4 1.47 mM à pH 7.4) et centrifugées à 4000 rpm une minute deux fois. Cette étape de lavage a été répétée à deux reprises pour éliminer toute trace de sérum présent dans le milieu de culture. Le culot de cellules a été resuspendu dans 10 à 60 μL de laemmli blanc 1.5x (150 mM Tris, 15% glycérol, 3.75% SDS) selon la grosseur du culot puis bouilli au moins 5 minutes à 95°C. L'échantillon protéique a été soniqué à une amplitude 90 pendant deux minutes avec 10 secondes de repos toutes les 20 secondes de sonication. Pour chaque échantillon, la quantité de protéines a été déterminée à partir d'une gamme étalon de BSA (protéine contrôle) réalisée au préalable avec du BCA à une DO de 562 nm.

Les échantillons protéiques ont migré dans un gel de polyacrylamide 12% (gel de résolution : Tris 0.38 M, pH 8.8, SDS 0.1% et acrylamide/bisacrylamide (30 : 0.8) 12% ; gel de concentration : Tris 0.13 M, pH 6.8, SDS 0.1% et acrylamide/bisacrylamide (30 : 0.8) 4%) dans le tampon de migration (Tris base 0.025 M, glycine 0.192 M et SDS 0.1%) à 100 V puis 160 V. La membrane de nitrocellulose a été utilisée pour le transfert des protéines dans le tampon de transfert (Tris base 0.01 M, glycine 0.077 M et éthanol 20%) pendant une heure à 115 V. Pour bloquer les sites non-spécifiques sur la membrane, celle-ci a été incubée une heure à température pièce avec du PBS tween-20 0.05 % et lait 5%. Toute la nuit (ou une heure), la membrane a été incubée à 4°C (ou à température ambiante) avec l'anticorps primaire. Après trois lavages au PBS tween-20 0.05 %, la membrane a été incubée une heure à température ambiante avec l'anticorps secondaire.

Grâce au mélange du substrat de la peroxydase HRP (le luminol) et la solution activatrice (peroxide) (50 : 50) (Millipore) déposé sur la membrane, l'enzyme liée à l'anticorps secondaire a produit de la lumière qui a été détecté sur des films.

18. Clonage

Le clonage permet d'intégrer une séquence d'ADN d'intérêt dans un plasmide d'expression.

Deux microgrammes du vecteur ont été digérés par les enzymes de restriction avec 1X du tampon compatible et 1/100 de l'albumine sérique bovine (BSA) durant 2 à 4 heures à 37°C. Les enzymes ont été inactivées à 65°C pendant 20 minutes. L'insert est isolé après digestion du plasmide d'origine par les mêmes enzymes qui ont digérés le vecteur ou amplifié par PCR avec des amorces qui contiennent les sites de restrictions des mêmes enzymes de digestion.

Le plasmide digéré et l'insert isolé ont été extraits du gel d'agarose après leur migration et coloration au bromure d'éthidium (BET). L'ADN a été extrait des morceaux d'agarose découpé grâce au kit de purification sur colonne « Qiagen PCR purification kit » (Cat. No. 28106). Le morceau d'agarose contenant l'ADN a été dissout dans 3 fois son volume de tampon QG (100 mg ~ 100 µL) à 50°C pendant 10 minutes. Un volume d'isopropanol a été ajouté au morceau d'agarose dissout. Pour lier l'ADN à la colonne, le mélange a été déposé sur une colonne, puis centrifugé pendant une minute à 13 000 rpm. Ces conditions de centrifugation ont été les mêmes tout le long de la purification. Pour enlever complètement l'agarose sur la colonne, 500 µL de tampon QG ont été déposé sur la colonne et centrifugé. La colonne a été lavée par 750 µL du tampon de lavage par centrifugation. Afin d'éliminer les traces d'éthanol, la colonne a été centrifugée à nouveau. Trente microlitre d'eau à pH 8 ont permis d'éluer les ADN d'intérêts pour effectuer les ligations et être conservés à -20°C.

Deux à trois ratios de ligation ont été testés en sachant que la quantité d'insert était toujours plus importante que celle du vecteur. Donc des ratios de 1/3 à 1/9 de vecteur (par rapport à l'insert) ont été ligué dans l'insert avec 1 μ L de la ligase T4 (Catherine Desrosiers) et 1X du tampon de la ligase T4 dans un volume final de 10 μ L. La ligation a été faite soit 1 heure à 37°C, soit 3 heures à température ambiante ou bien à 16°C toute la nuit. La ligase a été inactivée pendant 20 minutes à 65°C. La totalité du produit de ligation a été utilisé pour transformer les bactéries chimio-compétentes.

RESULTATS

CHAPITRE 1

Identification of *cis*-acting RNA elements that control the expression of the pro-apoptotic transcription factor TAF6 δ

Catherine Kamtchueng¹, Marie-Ève Stébenne, Aurélie Delannoy¹, Emmanuelle Wilhelm¹, Hélène Léger², Arndt Benecke^{3,4} and Brendan Bell¹

Avant-propos :

Cet article est l'aboutissement de plusieurs années de travail de Catherine KAMTCHUENG et de Marie-Eve STEBENNE (étudiante en maîtrise au laboratoire du Dr. BELL de 2006 à 2008). J'ai contribué à compléter les travaux de Marie-Eve STEBENNE sur le plan expérimental. J'ai participé entièrement à l'écriture de cet article avec mon directeur de recherche Dr. Brendan BELL. Cet article sera soumis en 2013.

Résumé de l'article :

La voie d'apoptose TAF6 δ peut dicter les décisions de vie ou de mort indépendamment du suppresseur de tumeur, p53. TAF6 δ est une sous-unité pro-apoptotique inductible du facteur général de transcription de l'ARN polymérase II (Pol II), TFIID. Le choix des sites d'épissage alternatif a été montré être un événement primordial déclenchant la mort par la voie de TAF6 δ . Rien n'est encore connu sur les mécanismes qui activent l'épissage de TAF6 δ . De plus, l'impact transcriptomique du gain de fonction de TAF6 δ versus la perte de fonction de la forme majeure d'épissage TAF6 α reste encore indéfini. Ici, nous montrons par une analyse de puce à ADN comparative que TAF6 δ conduit un profil transcriptomique distinct de celui de la déplétion de TAF6 α . Pour définir les éléments *cis* d'ARN régulant l'épissage alternatif de TAF6 δ , nous avons fait une analyse mutationnelle avec un système de minigène TAF6. Nos données mettent en évidence plusieurs nouveaux éléments d'ARN qui modulent TAF6 δ et révèlent un rôle de la structure secondaire d'ARN dans la sélection de TAF6 δ .

Identification of *cis*-acting RNA elements that control the expression of the pro-apoptotic transcription factor TAF6 δ

Catherine Kamtchueng¹, Marie-Ève Stébenne, Aurélie Delannoy¹, Emmanuelle Wilhelm¹, Hélène Léger², Arndt Benecke^{3,4} and Brendan Bell¹

1. RNA Group. Département de microbiologie et d'infectiologie, Faculté de médecine et sciences de la santé, Université de Sherbrooke, and Centre de recherche du CHUS, Pavillon de recherche appliquée sur le cancer, 3201 rue Jean-Migneault, Sherbrooke, Québec J1E 4K8, Canada.

2. Present Address: Fox Chase Cancer Center, Philadelphia, PA, USA; email: Helene.Leger@fccc.edu.

3. Institut des Hautes Etudes Scientifiques, Centre National de la Recherche Scientifique, Bures sur Yvette, France.

4. Present Address: Department of Microbiology, University of Washington, Seattle, USA; email: benecke@uw.edu

Abstract

The TAF6 δ pathway of apoptosis can dictate life versus death decisions independently of the status of p53 tumor suppressor. TAF6 δ is an inducible a pro-apoptotic subunit of the general RNA polymerase II (Pol II) transcription factor TFIID. Alternative splice site

choice of TAF6 δ has been shown to be a pivotal event in triggering death via the TAF6 δ pathway, yet the nothing is currently known about the mechanisms that promote TAF6 δ splicing. Furthermore the transcriptome impact of the gain of function of TAF6 δ function versus the loss of function of the major TAF6 α splice form remain undefined. Here we employ comparative microarray analysis to show that TAF6 δ drives a transcriptome profile distinct from that resulting from depletion of TAF6 α . To define the *cis*-acting RNA elements responsible for TAF6 δ alternative splicing we performed a mutational analysis of a TAF6 minigene system. The data point to several new RNA elements that can modulate TAF6 δ and also reveal a role for RNA secondary structure in the selection of TAF6 δ .

Introduction

The TAF6 δ pathway of apoptosis (Fig. 1) can control cell death versus life decisions of human cells {Bell, 2001 #327; Wilhelm, 2010 #979; Wilhelm, 2008 #976}. TAF6 δ is a splice variant of the TAF6 protein that is a core subunit of the general RNA polymerase II (Pol II) transcription factor, TFIID {Bieniossek, 2013 #1328; Wright, 2006 #858}. TFIID nucleates the formation of the Pol II pre-initiation complex and therefore represents a highly regulated step in the gene expression pathway of protein-coding genes {Burley, 1996 #39}. TFIID is the major core promoter recognition complex of the Pol II machinery and consists of TATA-binding protein (TBP) and a constellation of approximately 14 associated TBP-associated factors (TAFs) {Cler, 2009 #1341}. TAF6 δ is an inducible pro-apoptotic isoform of TAF6 that lacks 10 amino acids in its histone-fold domain. In contrast to the major TAF6 α isoform, TAF6 δ cannot interact with TAF9

and instead forms a TAF9-lacking complex termed TFIID π that drives a pro-apoptotic gene expression {Wilhelm, 2010 #979}. The TAF6 δ pathway has emerged as a model system to investigate the mechanisms by that transduce extracellular signals to trigger cellular suicide by impinging on the basal Pol II machinery. Moreover, because the TAF6 δ pathway induces cell death independently of p53 {Wilhelm, 2008 #976}, it represents a potential therapeutic target of strategic value for the killing tumor cells that frequently lack functional p53 {Vousden, 2007 #869}.

Modified antisense RNA oligonucleotides that force the splicing machinery to switch from producing a majority of TAF6 α to producing a majority of TAF6 δ in living cells trigger apoptosis demonstrating that changes in alternative splicing can trigger the TAF6 δ pathway of programmed cell death {Wilhelm, 2010 #979; Wilhelm, 2008 #976}. Alternative splicing plays a major role in proteomic diversification {Pan, 2008 #1342; Wang, 2008 #1343}. In the case of programmed cell death, alternative splicing can control cell life versus death decisions by regulating the balance of anti-apoptotic versus pro-apoptotic splice variants of genes within cell death pathways {Schwerk, 2005 #843}. *Cis*-acting RNA elements that can either enhance or silence the selection of alternative splice sites by the spliceosome to control splice site decisions {Chen, 2009 #1344}. These elements are classified upon their effect on a given splicing event and their location. Regulatory *cis*-acting RNA elements thus include exonic splicing enhancers (ESE), exonic splicing silencers (ESS), intronic splicing enhancers (ISE), and intronic splicing silencers (ISS). These *cis*-acting RNA elements act to recruit *trans*-acting protein factors, often from the SR protein family {Shepard, 2009 #1350} or the hnRNP family {Martinez-Contreras, 2007 #1351}. Layered upon the network of RNA-protein

interactions that underpin alternative splicing decisions is the key role of RNA secondary structure within the pre-mRNA that has impacts on splice site recognition by the spliceosome as well as on RNA-protein interactions {Buratti, 2004 #1340;Jin, 2011 #1337;McManus, 2011 #1338;Warf, 2010 #1339}.

One challenge in the study of alternative splicing is to define the relative biological impact the gain of the alternative splice form versus the loss of the constitutive form. While the induction of alternative splice variants often have important biological effects, in extreme cases alternative splicing serves only to dampen gene expression, as is the case when these events are coupled to the nonsense-mediated decay pathway {Lewis, 2003 #1349;Pan, 2006 #1348}. To shed light on the mechanism controlling the TAF6 δ pathway of apoptosis, here we have compared the transcriptome impacts of loss of function of the major TAF6 α splice variant via siRNA depletion versus those resulting from the induction of the pro-apoptotic TAF6 δ splice variant. The results reveal an essential function for TAF6 δ induction in the reprogramming of a specific pro-apoptotic transcriptome landscape. Despite the importance of inducible TAF6 δ expression, nothing is currently known about the mechanisms governing alternative TAF6 δ splicing. We therefore developed and validated a minigene system for the mutational dissection of TAF6 *cis*-acting RNA elements. We report here the first identification of RNA elements that can influence splicing of TAF6 δ .

Materials and Methods

Cell culture

HeLa cell line was maintained in culture in Dulbecco's modified Eagle's medium supplemented with 2.5% fetal calf serum and 2.5% calf serum.

Plasmids

To construct the TAF6 minigene (pTAF6mg), the genomic region of TAF6 containing exon 2 to exon 3 was amplified by PCR from HeLa cell genomic DNA with primers 5'-AAAAAGGGATCCCATGGGCATCGCCCAGATTCAGG-3' (forward) and 5'-AAAAAGGAATTCCAAGGCGTAGTCAATGTCACTGG-3' (reverse). The PCR product was ligated into pTZ57R/T (Fermentas). The new plasmid was digested with EcoRI and BamHI and the TAF6 fragment was inserted into the same sites of pcDNA3.1+. The mutated minigenes were created by PCR mutagenesis using Pfu DNA polymerase with specific primers bearing mutations {Papworth, 1996 #1214} (all sequences of oligonucleotides used in this study are available upon request).

Transfections

siRNA 5'-rGrGrArGrUrGrUrCrCrArGrArArGrUrArCrArUrCrGrUrGGT-3' (T6-1) and 5'-rGrGrArGrUrGrUrCrCrArGrArArGrUrArCrArUrCrGrUrGGT-3' (T6-2) were

employed to deplete all known splice variants of TAF6. SiRNA were transfected at a final concentration of 10nM with lipofectamine 2000 (Invitrogen) as a delivery agent (1.6 μ l/ml) according to the manufacturer's recommendations. 250 to 300ng of wild-type or mutant TAF6 minigene were transfected with 1 μ l DMRIE-C (Invitrogen) per well in 24 well plate according to the manufacturer's recommendations. Cells were transfected by siRNA T6-1 two times. The second transfection was performed 24h after the first one and culture was maintained for a total of 48 or 72h before harvesting cells for RNA extraction (RNeasy Qiagen) for microarray analysis or 64h for protein analysis. All transfections were performed with OptiMEM medium (Invitrogen). Each transfection experiment was repeated three times.

Viability assay by methylene blue staining

75000 HeLa cells were split in 24 well plates and transfected 12 hours later with 10 nM dsRNA combined with lipofectamine 2000 (Invitrogen) as recommended by the supplier. DsiRNA were transfected again 24 and 48h after the first transfection. The culture was maintained for a total of 4 days after the first transfection. The culture medium was removed and the cell monolayer was washed carefully with 500 μ l PBS. Cells were stained for 30 min at RT by the addition of 200 μ l of a solution containing 5mg/ml of methylene blue in 50% ethanol. The plate was carefully and extensively washed with water until no blue stain remained in the water. The plate was air dried completely. 500 μ l of a PBS solution containing 10mg/ml N-lauroyl sarcosine (Sigma L-5125) was added to each well. Lysis was performed for 1h at RT. 100 μ l of each lysate was used to measure

absorbance at 595 nm ($A_{515\text{nm}}$ =control), corresponding to methylene blue incorporation and cell content.

RT-PCR

Total RNA was extracted from cells using Trizol (Invitrogen) according to the manufacturer's recommendations. RNA was treated with 1 unit of DNase I (Promega) for 30 minutes at 37°C to remove any contaminating DNA. 1 µg of total RNA was reverse transcribed using MMLuV reverse transcriptase. Specific oligonucleotides for endogenous TAF6 (forward 5'-ATGGGCATCGCCCAGATTCAGG-3' and reverse 5'-AAGGCGTAGTCAATGTCACTGG-3') and exogenous TAF6 minigene constructs (forward 5'-ATGGGCATCGCCCAGATTCAGG-3' and reverse 5'-AATAGCGATCCACGCGACTAGTGG-3') were used for PCR amplifications of 1/10 of the total cDNA (25 cycles, 1 min at 94°C, 45 sec at 58°C, 50 sec at 68°C, initial step 3 min at 95°C, final extension 5 min at 68°C), using Taq DNA polymerase. The splicing isoforms were quantified by capillarity electrophoresis on a BioAnalyser 2100 (Agilent) according to the manufacturer's instructions.

Western blot analysis

After treatment as described, cells were lysed in 1.5 X laemmli sample buffer, sonicated and electrophoresed on 7.5 and 12% SDS-PAGE before electro-transfer to PVDF membrane. Blots were probed with previously described antibodies against TAF6α,

TAF6 δ or TBP {Bell, 2001 #327} followed by goat anti-mouse HRP-conjugate secondary antibodies (Jackson ImmunoResearch Laboratories) before enhanced chemiluminescence detection.

Microarrays

RNA isolation and quality control were performed as previously described {Wilhelm, 2010 #979; Wilhelm, 2008 #987}. Applied Biosystems HGS V2 arrays {Noth, 2006 #895} were hybridized, washed, and exposed according to the protocols by the technology provider. Raw data were quality controlled {Brysbaert, 2010 #1322}, and normalized using NeONORM (k-parameter=0.2) {Noth, 2006 #897}. Longitudinal analysis was performed using a Kohonen-maps based classifier as described in {Rasmussen, 2012 #1321}, using the CDS statistical test {Tchitchek, 2012 #1320}. The data are freely accessible through the MACE database at <http://mace.ihes.fr> under accession no.: 2418870990.

Results

SiRNA-mediated depletion of TAF6 causes a loss of viability in human cells

We have previously analyzed the transcriptome impact of the induction of TAF6 δ by using microarray experiments {Wilhelm, 2010 #1361}. To define the impact of the loss of total TAF6 protein in human cells we performed transcriptome-wide analysis of gene expression following depletion of total TAF6 proteins by siRNA. Importantly, the minor

TAF6 δ protein isoform is not expressed at detectable levels in HeLa cells under normal growth conditions, so that the effects of the siRNA cannot be due to significant reductions in TAF6 δ protein levels {Wilhelm, 2010 #979; Wilhelm, 2008 #976}. siRNAs were designed and validated for their ability to reduce TAF6 protein levels. Two siRNA resulted in a 60% to 90% reduction in TAF6 protein (Fig. 2A and 2B). To determine the impact of the loss of total TAF6 protein on cell viability we employed the methylene blue colorimetric assay (see materials and methods). Both siRNAs that depleted TAF6 resulted in a concomitant loss of cell viability (Fig. 2C). We conclude that TAF6 expression is essential for cell viability in human cells, extending previous results demonstrating that TAF6 is essential for viability in distinct organisms including *S. cerevisiae* {Michel, 1998 #173; Shen, 2003 #490}, *Drosophila* {Aoyagi, 2001 #884} and zebrafish {Amsterdam, 2004 #880}. We next used a siRNA that reduces total TAF6 α expression and results in the loss of viability to further determine the impact of loss of TAF6 α on the transcriptome using microarray analysis. Total RNA was isolated from HeLa cells treated with siRNA directed against TAF6 and from cells treated with control siRNA for microarray analysis {Noth, 2005 #898; Wilhelm, 2010 #979; Wilhelm, 2008 #987}. Two time points (48 and 72 hours) were chosen for transcriptome analysis to ensure the detection of the broadest possible number of TAF6-dependent transcripts as well as to ensure measurements before the onset of massive cell death. The levels of TAF6 mRNA were internally controlled in the transcriptome analysis by probes for *taf6* and showed a reduction to greater than 80% at both 48 and 72 hours post-transfection (Fig. 2D).

Distinct transcriptome impacts of depleting TAF6 α versus inducing TAF6 δ

The knock-down of total TAF6 resulted in statistically significant changes in gene expression levels including a global reduction in transcription accompanied by the increase in a minority of mRNA transcripts 48 hours after siRNA transfection and particularly after 72 hours (Fig. 3A). We also compared the changes in gene expression of the 961 previously identified TAF6 δ -dependent transcripts {Wilhelm, 2010 #979} to changes in response to siRNA depletion of total TAF6 and found that the two conditions result in highly distinct transcriptome landscapes (Fig. 3C). To examine more closely the changes resulting from the induction of TAF6 δ versus depletion of TAF6, we filtered the microarray data to compare the overlap between the two sets of regulated transcripts. Of 961 TAF6 δ -dependent transcripts 128 were also significantly regulated by the siRNA against TAF6 at either 48 or 72 hours (Fig. 3D). Of the 128 regulated the majority (81) were oppositely regulated (induction versus repression). 44 transcripts were induced by both the loss of TAF6 and the induction of TAF6 δ , and the overlap between these genes was statistically significant ($P = 7.86 \times 10^{-16}$, hypergeometric distribution). In contrast, 1513 genes were repressed in the absence of TAF6 and only three transcripts were repressed in both conditions (Fig. 3D) and there was no statistically significant overlap between repressed genes. The microarray results revealed a minor subset of common effects in the induction of TAF6 δ versus depletion of total TAF6 transcripts. Importantly, the data also reveal that the induction of TAF6 δ drives a distinct pro-apoptotic gene expression program and therefore underscore the necessity for TAF6 δ isoform expression to trigger this specific pathway of apoptosis. To further explore the effects of TAF6 depletion gene ontology analysis was applied. Of the seven pathways that are statistically

overrepresented during TAF6 depletion, only three including the integrin signaling pathway, the p53 pathway and apoptosis signaling pathways were also found in the TAF6 δ transcriptome signature (Fig. 3E) {Wilhelm, 2010 #979}. We also noted that the morphology of cells dying due to siRNA depletion of TAF6 did not exhibit membrane blebbing by light microscopy (Fig. S1). This morphology was distinct from that observed in TAF6 δ -induced apoptosis where classical membrane blebbing is readily observed {Wilhelm, 2010 #979}, suggesting that TAF6 δ -induced death and death by depletion of TAF6 are biochemically distinct forms of death. Taken together the microarray data reveals that the loss of total TAF6 has a highly distinct impact on cellular gene expression programs compared to those resulting from increases in levels of the TAF6 δ splice variant.

Development and validation of a minigene system to study TAF6 alternative splicing

Having established the importance of TAF6 δ induction versus the loss of TAF6 α in driving a specific transcriptome profile and cell death programme, we next sought to investigate the molecular mechanisms that control the expression of the minor TAF6 δ splice variant versus the major TAF6 α splice variant. We cloned the genomic region from the TAF6 gene that contains part of exon 2, the alternative portion of exon 2, the natural full-length intron 2, and a part of exon 3 into the eukaryotic expression pcDNA3.1+ under the control of the CMV promoter (Fig. 4A). HeLa cells were chosen as a model to study TAF6 alternative splicing since the pro-apoptotic TAF6 δ splice variant was originally cloned from a HeLa cell cDNA library {Bell, 2001 #327}. PCR primers specific to the transcribed flanking vector sequences were used to amplify

exogenous TAF6 transcripts after transient transfection of the minigene into HeLa cells (Fig. 4A). PCR products corresponding to the expected sizes for the minigene products of unspliced pre-messenger, the major TAF6 α transcript, and the minor TAF6 δ transcript were detected by RT-PCR (Fig. 4A). Sequencing of each of these PCR products showed that they corresponded exactly to the expected splice products (B.B., unpublished results). Importantly, the ratio of TAF6 δ versus total transcripts was found to range from 10% to 20% between independent transfections (Figs. 4B, 5A and 5B), a percentage that corresponds well to the endogenous TAF6 splicing pattern in HeLa cells {Wilhelm, 2008 #976}. The growth conditions of independent transfections appeared to influence basal levels of TAF6 δ since biological triplicates within a single wave of transfections were highly reproducible (Figs. 4-8). Taken together, these data indicate that the plasmid minigene system accurately recapitulates endogenous TAF6 δ splicing patterns in transfected HeLa cells.

To further characterize the minigene system we tested constructs bearing point mutations within the one of the two alternative 5' splice sites (SSs). When we crippled the proximal (α) 5' SS by mutating the first two nucleotides of the intronic SS from the consensus GT to CA, the transfected minigene spliced uniquely at the alternative distal (δ) 5' SS, as expected (Fig. 4B, Pko). Likewise, the same crippling mutation of the distal (δ) 5' SS resulted in splicing in HeLa cells uniquely at the proximal (α) 5' SS (Fig. 4B, Dko). We used a web tool http://rulai.cshl.edu/new_alt_exon_db2/HTML/score.html provided by Zhang laboratory (Cold Spring Harbor Laboratories) to calculate the alternative distal 5' SS and the constitutive proximal 5' SS strength scores as 2.8 and 6.4 respectively. These are relatively weak given that a perfect consensus site score is 12.6

and the average value for a constitutive 5' SS is 8.1. To test the impact of splice site strength we mutated both sites individually to the consensus (AG/GUAAGU). When the proximal 5' SS is changed to the consensus no residual detection of the alternative δ variant was detected (Fig. 4B, Pcons). When the alternative distal 5' SS was changed to a consensus sequence all of the splice products used the distal site (Fig. 4B, Dcons). We conclude that one parameter that impacts the alternative splice site choice of TAF6 is the complementarity of the 5' SSs to U1 snRNP as expected. We further conclude that the minigene system we developed provides a useful system to dissect the *cis*-acting RNA elements that control the expression of TAF6 δ .

Having established a minigene system to study TAF6 alternative splicing we next set out to perform a mutational dissection of the RNA elements within the minigene that impact the expression of the pro-apoptotic TAF6 δ isoform. We have applied a combination of resources to guide our mutational analysis including predictive algorithms for RNA binding proteins and *cis*-acting elements including RESCUE-ESE {Fairbrother, 2004 #1324}, ESE finder {Cartegni, 2003 #678}, and "Splicing Rainbow" {Stamm, 2006 #1323}. We also considered evolutionary conservation of intronic sequences and employed unbiased scanning mutagenesis. For simplicity we have subdivided the minigene into several sub-regions and our mutational analysis of each region is presented below.

Scanning mutagenesis of constitutive exon 2 and exon 3

To search for *cis*-acting RNA elements we initially performed scanning by mutating blocks of 10 nucleotides within the constitutive portion of exon 2 (Fig. 5A). The mutation

of two nucleotide blocks Exon 2-1 and Exon 2-3 had no significant effect on the TAF6 splicing pattern (Fig. 5A). Another nucleotide block in exon 2, exon 2-2 produced reproducible reductions in the levels of TAF6 δ (Fig. 5A). These results suggest a *cis*-acting RNA sequence within the exon 2-2 can favour the selection of the distal (δ) 5' SS. We note that in HeLa cells the endogenous basal levels (~10%) of TAF6 δ do not produce detectable TAF6 δ protein by Western blotting or immunofluorescence with sensitive antibodies, suggesting the exon 2-2 mutation would not be expected to have profound biological impacts. Nevertheless, when measured at the mRNA level the mutation of the exonic element 2-2 can reduce the splice site choice of TAF6 δ .

We next used the same scanning mutagenesis to query whether exon 3 contained RNA elements important for the selection of TAF6 δ . We mutated three 10 nucleotide blocks near to the 3' SS of our TAF6 minigene (Fig 5B). We found no significant differences in the ratio of TAF6 δ for any of the three mutations exon 3-1, 3-2 or 3-3 (Fig. 5B). We conclude that the region of exon 3 proximal to intron 2 does not play a major role in the selection of the TAF6 δ splice variant.

Targeted mutagenesis of intron 2 of the *taf6* gene

To begin characterizing intron 2 of the minigene, we first examined sequence homology between the ninety-nine nucleotide long natural human intron 2 and forty-six vertebrate species (<http://genome.ucsc.edu/>). Minimal informative sequence conservation was found within intron 2, however a small highly conserved motif that fits the highly degenerate yUnAy consensus {Gao, 2008 #1329} for a human branch point site was identified at nucleotides 153 - 157 (Fig. S2). To determine if the adenine at position 156 corresponds

to the branchpoint we mutated it to a guanosine. Upon transfection in HeLa cells the mutated minigene showed profoundly reduced splicing activity (Fig. S3). The residual splicing activity is low but detectable and could represent the inefficient use of alternative neighboring adenines. We conclude that adenine 156 is essential for splicing of the TAF6 minigene and most likely represents the major branchpoint site in intron 2.

To further explore *cis*-acting RNA elements within intron 2 that could impact the selection of the distal 5' SS, we chose candidate motifs after manual and software-assisted (see above) analysis of potential regulatory sequence motifs. We decided to test the roles of several sequence motifs in the TAF6 minigene system including poly C motifs {Expert-Bezancon, 2002 #1334}, poly G motifs {Sirand-Pugnet, 1995 #1333} and a potential binding site for hnRNPA1 {Yang, 1994 #1335} because these motifs have been previously shown to influence alternative splicing. The mutation of a potential hnRNPA1 binding site {Ishikawa, 1993 #1336} within intron had no significant effect on the TAF6 splicing pattern (Fig. 6A, koA1). Likewise, a mutation of the CCCC (C₄) motif at position 175-178 had little effect on TAF6 alternative splicing (Fig. 6A, C₄-175). The mutation of a C₃ motif at positions 125-127 and C₄ combined significantly reduced levels of the TAF6 δ splice variant (Fig. 6A, C₃-125; C₄-132). To determine if one of these motifs was more important for TAF6 δ expression we mutated them individually and found that mutation of C₃-125 alone caused a small but measurable reduction in TAF6 δ levels (Fig. 6A, C₃-125). Mutation of the C₄-132 motif alone caused a reduction in TAF6 δ expression similar to that of the double mutation (Fig. 6B, C₄-132), suggesting that the C₄-132 motif plays the predominant role in TAF6 δ splice site selection. To complete our investigation of poly C motifs, we mutated a C₃ motif at position 91-93 in

the minigene and found it produced modestly reduced splice selection of TAF6 δ (Fig. 6A, C₃-91).

We next turned our attention to the potential function of three poly G sequences within intron 2 of the TAF6 minigene. The mutation of a G₃ motif position 113-115 resulted in reduced expression of the TAF6 δ splice variant (Fig. 6B, G₃-113). The mutation of a G₃ motif position 129-131 had little impact on TAF6 splicing ratios (Fig. 6B, G₃-129). The mutation of a G₃ motif position 150-152 generated wild type splicing ratios. (Fig. 6B, G₃-150). Taken together, our analysis of intron 2 of the TAF6 minigene define a putative branchpoint adenosine at position 156 that is essential for splicing, and shows that specific poly C and poly G motifs in the intron can enhance the selection of the pro-apoptotic TAF6 δ splice form.

RNA secondary structure at the proximal 5' splice site can force the selection of TAF6 δ

We then focused on the alternative exon 2 (exon 2a) a critical region of the minigene because it lies physically between the α and δ alternative 5' splice sites (Fig. 4A), and because modified antisense RNA oligonucleotides that anneal to it can shift the splicing from the major to the pro-apoptotic δ form in living cells {Wilhelm, 2008 #976}. The ESEfinder algorithm {Cartegni, 2003 #678} was employed and detected a potential SRSF1 binding site in exon 2a (Fig. 7A). To test a potential role for the SRSF1 motif in the 5' splice site choice we designed two point mutations to prevent (Fig. 7A, SRSF1-) or enhance (Fig. 7A, SRSF1+) SRSF1 binding. Upon transfection these mutations both gave strong increases in the levels of the δ splice form (Fig 7A), a result not compatible with a

role for SRSF1 binding to this motif. To further dissect the impact of the two-nucleotide mutation within SRSF1+ that caused a complete reversal of the splicing pattern, we mutated these positions individually. Mutation of position 68 alone had little effect on the splicing pattern (Fig. 7A, A68C). In contrast, mutation of adenosine 73 to guanosine alone resulted in a complete shift towards the usage of the distal δ 5' SS (Fig. 7A, A73G).

Having ruled out a role for SRSF1, we sought an alternative hypothesis for the strong impact of this mutation. Given the well-documented importance of RNA secondary structure on alternative splicing {Buratti, 2004 #1340; Jin, 2011 #1337; McManus, 2011 #1338; Warf, 2010 #1339}, we employed the M-Fold algorithm {Zuker, 2003 #1330} to test for potential impacts on RNA structure. When compared to the wild type sequence the A73G mutation was predicted to form a stem-loop that based on precedence {Hutton, 1998 #475} would be stable enough to potentially block U1 snRNP from binding to the proximal 5' splice site (Fig. 7A, at right). To test the hypothesis that mutation A73G reverses TAF6 splicing via the formation of secondary structure we performed further mutagenesis to provide support for the putative G73 - C86 base pair in cells. Mutation of C86U in the wild type context A73 did not change TAF6 splicing pattern (Fig. 7B, mutant AU). We also performed mutation CG that is predicted to form a stem-loop of equal stability to A73G replacing the natural C86 within position 5 of the proximal 5' splice site with G to increase its strength. This construct showed less δ splicing than the wild type (Fig. 7 B mutant CG), compatible with a competition between splice site strength and RNA secondary structure in cells. To probe the weak potential secondary structure formed by the wild type structure could impact TAF6 splicing we performed a series of mutations to weaken such a structure, but found no strict correlation

between the stability of such a structure and splice site selection (Fig. 7C). To further confirm the putative structure formed by the A73G mutation we performed mutations that disrupt the structure while leaving the proximal 5' SS untouched and found no change in splicing (Fig. 7D, GC 74). Mutations that weaken but do not ablate the putative A73G-induced stem-loop showed intermediate levels of A point mutation in the loop sequence had no effect, as expected (Fig. 7D, GC 78). Taken collectively, these findings indicate that a stem-loop structure competing for the proximal 5' SS can strongly enhance the use of the distal δ 5' SS.

Evidence for an exonic splicing silencer in alternative exon 2

Given that local secondary structure occurs at the proximal 5' SS, and that it is located only 30 nucleotides from the distal 5' SS, we postulated that competition between RNA structure and the distal 5' SS might also play a role in the expression of TAF6 δ . We used the M-Fold algorithm {Zuker, 2003 #1330} to predict potential secondary structures overlapping the distal 5' SS. The most probable theoretical stem-loop structure predicted could potentially occlude the interaction of U1 snRNP interaction with the distal δ 5' SS (Fig. 8A). To provide evidence to confirm or exclude the existence of such a stem-loop we constructed a series of minigenes with mutations that would weaken, destroy or maintain it. No correlation was observed between potential stem-loop strength and the splicing outcome, ruling out a major role for such a structure in the regulation of TAF6 alternative splicing (Fig. 8A).

Interestingly, however, our mutational analysis revealed that a single nucleotide change (C66U) produced a greater than two-fold increase in the production of the TAF6 δ

splice form (Fig. 8A and B, C66U). To further delineate the role of cytosine 66 we mutated it to uridine, adenosine and guanosine respectively. Mutations C66U and C66A both resulted in clear increases in TAF6 δ splice form production (Fig. 8C), even though these nucleotides have distinct hydrogen bonding specificities within RNA secondary structures. C66G had much less effect than these mutations, although the mutation would also be expected to have alter base-pairing within RNA structure. Based on the above results we conclude that cytosine 66 is part of an exonic splicing silencer (ESS) that most likely acts through *trans*-acting factors rather than RNA secondary structure. The mutational data described above support a hypothetical model for *cis*-acting RNA elements in the regulation of TAF6 alternative splicing that is presented schematically in Figure 9.

Discussion

TAF6 a core subunit of the general RNA Pol II transcription factor TFIID {Wright, 2006 #858}, and has previously demonstrated to be essential for viability in several organisms including the budding yeast *Saccharomyces cerevisiae* {Michel, 1998 #173; Poon, 1995 #190}, *Drosophila* fruit flies {Aoyagi, 2001 #884}, the flowering plant *Arabidopsis thaliana* as well as the fresh water fish *Danio rerio* {Amsterdam, 2004 #880}. Here we show for the first time that TAF6 is essential for viability in human cells (Fig. 2). Although this result is not unexpected, the significance of the finding derives from the potential of TAF6 as a therapeutic target for numerous diseases that result from deregulated apoptosis {Reed, 2002 #611}. Many apoptotic genes are required for normal development at the organismal level but tumor suppressors such as BRCA1, PTEN,

CDKN2A (ARF/p16INK4), RB1, APC, and p53, Bcl-2 family members, the caspases and death receptors are all dispensable for viability at the cellular level {Ranger, 2001 #990; Vogelstein, 2004 #989}. TAF6 therefore represents a rare class of genes that are essential for cell viability, but also possess splice variants with potent pro-apoptotic activity. As an essential gene with pro-apoptotic potential TAF6 it is of high strategic interest in the development of anti-cancer treatments that avoid the development of chemoresistance and to target p53 negative tumors {Watson, 2013 #1352}.

Using a minigene system that recapitulates the endogenous TAF6 splicing pattern we identified several *cis*-RNA elements that modulate TAF6 δ splicing as schematically shown in Figure 9. The sensitivity of the proximal 5' splice site evokes the possibility that the modulation of RNA folding could contribute to the physiological selection of TAF6 δ . Further work will be required to confirm or exclude this hypothesis. The first mapping of the *cis*-acting RNA within elements provides the essential groundwork for future studies to identify *trans*-acting factors that regulate TAF6 δ splicing. Indeed, the identification of the important RNA elements will be crucial for both proteomic {Kar, 2011 #1355} and genomic {Moore, 2010 #1353} approaches to identify *trans*-acting splice regulatory proteins. A limitation of the current study is the fact that the mutations could conceivably differentially alter RNA stability in addition to splice site choice. In addition potential long-range *cis*-acting RNA elements will not have been identified due to the limited size of our TAF6 minigene. Further work will be required to address these possibilities.

Accumulating evidence points to a potential link between the TAF6 δ pathway and cancer biology. A cDNA encoding the major TAF6 α variant was identified in a large-scale screen as being able to increase colony formation in human hepatocellular

carcinoma cells and mouse embryonic fibroblasts {Wan, 2004 #991}. Integrative genomic data show that the *taf6* gene is amplified in lung cancer {Campbell, 2008 #992}. TAF6 has been reported as a genomic marker of poor prognosis in lung adenocarcinoma {Aviel-Ronen, 2008 #1009}. TAF6 mRNA was identified as being overexpressed in inflammatory breast cancer {Dressman, 2006 #993}. Interestingly, a specific splice variant of TAF6 with an extended exon 2 is reportedly overrepresented in ductal cell carcinoma {Wang, 2004 #863}. Taken together these findings suggest that the major anti-apoptotic TAF6 α splice variant possesses oncogenic potential. In stark contrast, the minor TAF6 δ splice variant has pro-apoptotic activity evoking a potential tumor suppressor activity {Bell, 2001 #327; Wilhelm, 2010 #979; Wilhelm, 2008 #976}. It is conceivable that anti-apoptotic TAF6 α expression can be decoupled from pro-apoptotic TAF6 δ in certain tumor types. The mapping of key *cis*-acting RNA elements we present here paves the way to experimentally test the existence of mutations in the essential *taf6* gene that could reduce or prevent TAF6 δ expression in human tumors.

Acknowledgments

We thank Dr. Juan Valcárcel for insightful discussions and support in the initial stages of this work. We thank Alexandre Magdzinski for technical support. We gratefully acknowledge Annie Leclerc for sharing her expertise in artwork. B. Bell is a member of the FRSQ-funded Centre de recherche clinique Étienne-Le Bel.

Author contributions

Conceived and designed experiments: BB and AB. Performed experiments: CK, MES, AD, EW, HL. Analyzed data HL, AB and BB. Wrote the manuscript: CK and BB.

Funding:

This work was supported through a Natural Science and Engineering Research Council of Canada Discovery Grant awarded to BB.

Competing Interests:

The authors declare that no competing interests exist.

References

1. Bell B, Scheer E, Tora L (2001) Identification of hTAF(II)80 delta links apoptotic signaling pathways to transcription factor TFIID function. *Mol Cell* 8: 591-600.
2. Wilhelm E, Kornete M, Targat B, Vigneault-Edwards J, Frontini M, et al. (2010) TAF6delta orchestrates an apoptotic transcriptome landscape and interacts functionally with p53. *BMC Molecular Biology* 11: 1.
3. Wilhelm E, Pellay FX, Benecke A, Bell B (2008) TAF6delta controls apoptosis and gene expression in the absence of p53. *PLoS ONE* 3: e2721.
4. Bieniossek C, Papai G, Schaffitzel C, Garzoni F, Chaillet M, et al. (2013) The architecture of human general transcription factor TFIID core complex. *Nature*.
5. Wright KJ, Marr MT, 2nd, Tjian R (2006) TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. *Proc Natl Acad Sci U S A* 103: 12347-12352.
6. Burley SK, Roeder RG (1996) Biochemistry and structural biology of transcription factor IID (TFIID). *Annu Rev Biochem* 65: 769-799.

7. Cler E, Papai G, Schultz P, Davidson I (2009) Recent advances in understanding the structure and function of general transcription factor TFIID. *Cellular and molecular life sciences* : CMLS 66: 2123-2134.
8. Vousden KH, Lane DP (2007) p53 in health and disease. *Nat Rev Mol Cell Biol* 8: 275-283.
9. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature genetics* 40: 1413-1415.
10. Wang ET, Sandberg R, Luo S, Khrebukova I, Zhang L, et al. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* 456: 470-476.
11. Schwerk C, Schulze-Osthoff K (2005) Regulation of apoptosis by alternative pre-mRNA splicing. *Mol Cell* 19: 1-13.
12. Chen M, Manley JL (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nature reviews Molecular cell biology* 10: 741-754.
13. Shepard PJ, Hertel KJ (2009) The SR protein family. *Genome biology* 10: 242.
14. Martinez-Contreras R, Cloutier P, Shkreta L, Fiset JF, Revil T, et al. (2007) hnRNP proteins and splicing control. *Advances in experimental medicine and biology* 623: 123-147.
15. Buratti E, Baralle FE (2004) Influence of RNA secondary structure on the pre-mRNA splicing process. *Molecular and cellular biology* 24: 10505-10514.
16. Jin Y, Yang Y, Zhang P (2011) New insights into RNA secondary structure in the alternative splicing of pre-mRNAs. *RNA biology* 8: 450-457.
17. McManus CJ, Graveley BR (2011) RNA structure and the mechanisms of alternative splicing. *Current opinion in genetics & development* 21: 373-379.
18. Warf MB, Berglund JA (2010) Role of RNA structure in regulating pre-mRNA splicing. *Trends in biochemical sciences* 35: 169-178.
19. Lewis BP, Green RE, Brenner SE (2003) Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proceedings*

- of the National Academy of Sciences of the United States of America 100: 189-192.
20. Pan Q, Saltzman AL, Kim YK, Misquitta C, Shai O, et al. (2006) Quantitative microarray profiling provides evidence against widespread coupling of alternative splicing with nonsense-mediated mRNA decay to control gene expression. *Genes & development* 20: 153-158.
 21. Papworth C, Bauer JC, Braman J, Wright DA (1996) Site-directed mutagenesis in one day with >80% efficiency. *Strategies* 9: 3-4.
 22. Wilhelm E, Pella FX, Benecke A, Bell B (2008) Determining the impact of alternative splicing events on transcriptome dynamics. *BMC Res Notes* 1: 94.
 23. Noth S, Brysbaert G, Pella FX, Benecke A (2006) High-sensitivity transcriptome data structure and implications for analysis and biologic interpretation. *Genomics Proteomics Bioinformatics* 4: 212-229.
 24. Brysbaert G, Pella FX, Noth S, Benecke A (2010) Quality assessment of transcriptome data using intrinsic statistical properties. *Genomics, proteomics & bioinformatics* 8: 57-71.
 25. Noth S, Brysbaert G, Benecke A (2006) Normalization using weighted negative second order exponential error functions (NeONORM) provides robustness against asymmetries in comparative transcriptome profiles and avoids false calls. *Genomics Proteomics Bioinformatics* 4: 90-109.
 26. Rasmussen AL, Tchitchek N, Susnow NJ, Krasnoselsky AL, Diamond DL, et al. (2012) Early transcriptional programming links progression to hepatitis C virus-induced severe liver disease in transplant patients. *Hepatology* 56: 17-27.
 27. Tchitchek N, Dzib JF, Targat B, Noth S, Benecke A, et al. (2012) CDS: a fold-change based statistical test for concomitant identification of distinctness and similarity in gene expression analysis. *Genomics, proteomics & bioinformatics* 10: 127-135.
 28. Michel B, Komarnitsky P, Buratowski S (1998) Histone-like TAFs are essential for transcription in vivo. *Mol Cell* 2: 663-673.

29. Shen WC, Bhaumik SR, Causton HC, Simon I, Zhu X, et al. (2003) Systematic analysis of essential yeast TAFs in genome-wide transcription and preinitiation complex assembly. *Embo J* 22: 3395-3402.
30. Aoyagi N, Wassarman DA (2001) Developmental and transcriptional consequences of mutations in *Drosophila* TAF(II)60. *Mol Cell Biol* 21: 6808-6819.
31. Amsterdam A, Nissen RM, Sun Z, Swindell EC, Farrington S, et al. (2004) Identification of 315 genes essential for early zebrafish development. *Proc Natl Acad Sci U S A* 101: 12792-12797.
32. Noth S, Benecke A (2005) Avoiding inconsistencies over time and tracking difficulties in Applied Biosystems AB1700/Panther probe-to-gene annotations. *BMC Bioinformatics* 6: 307.
33. Fairbrother WG, Yeo GW, Yeh R, Goldstein P, Mawson M, et al. (2004) RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. *Nucleic acids research* 32: W187-190.
34. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR (2003) ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res* 31: 3568-3571.
35. Stamm S, Riethoven JJ, Le Texier V, Gopalakrishnan C, Kumanduri V, et al. (2006) ASD: a bioinformatics resource on alternative splicing. *Nucleic acids research* 34: D46-55.
36. Gao K, Masuda A, Matsuura T, Ohno K (2008) Human branch point consensus sequence is yUnAy. *Nucleic acids research* 36: 2257-2267.
37. Expert-Bezancon A, Le Caer JP, Marie J (2002) Heterogeneous nuclear ribonucleoprotein (hnRNP) K is a component of an intronic splicing enhancer complex that activates the splicing of the alternative exon 6A from chicken beta-tropomyosin pre-mRNA. *The Journal of biological chemistry* 277: 16614-16623.
38. Sirand-Pugnet P, Durosay P, Brody E, Marie J (1995) An intronic (A/U)GGG repeat enhances the splicing of an alternative intron of the chicken beta-tropomyosin pre-mRNA. *Nucleic acids research* 23: 3501-3507.
39. Yang X, Bani MR, Lu SJ, Rowan S, Ben-David Y, et al. (1994) The A1 and A1B proteins of heterogeneous nuclear ribonucleoparticles modulate 5' splice site

- selection in vivo. *Proceedings of the National Academy of Sciences of the United States of America* 91: 6924-6928.
40. Ishikawa F, Matunis MJ, Dreyfuss G, Cech TR (1993) Nuclear proteins that bind the pre-mRNA 3' splice site sequence r(UUAG/G) and the human telomeric DNA sequence d(TTAGGG)_n. *Molecular and cellular biology* 13: 4301-4310.
 41. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research* 31: 3406-3415.
 42. Hutton M, Lendon CL, Rizzu P, Baker M, Froelich S, et al. (1998) Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature* 393: 702-705.
 43. Poon D, Bai Y, Campbell AM, Bjorklund S, Kim YJ, et al. (1995) Identification and characterization of a TFIID-like multiprotein complex from *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 92: 8224-8228.
 44. Reed JC (2002) Apoptosis-based therapies. *Nat Rev Drug Discov* 1: 111-121.
 45. Ranger AM, Malynn BA, Korsmeyer SJ (2001) Mouse models of cell death. *Nat Genet* 28: 113-118.
 46. Vogelstein B, Kinzler KW (2004) Cancer genes and the pathways they control. *Nat Med* 10: 789-799.
 47. Watson J (2013) Oxidants, antioxidants and the current incurability of metastatic cancers. *Open biology* 3: 120144.
 48. Kar A, Fushimi K, Zhou X, Ray P, Shi C, et al. (2011) RNA helicase p68 (DDX5) regulates tau exon 10 splicing by modulating a stem-loop structure at the 5' splice site. *Molecular and cellular biology* 31: 1812-1821.
 49. Moore MJ, Wang Q, Kennedy CJ, Silver PA (2010) An alternative splicing network links cell-cycle control to apoptosis. *Cell* 142: 625-636.
 50. Wan D, Gong Y, Qin W, Zhang P, Li J, et al. (2004) Large-scale cDNA transfection screening for genes related to cancer development and progression. *Proc Natl Acad Sci U S A* 101: 15724-15729.

51. Campbell JM, Lockwood WW, Buys TP, Chari R, Coe BP, et al. (2008) Integrative genomic and gene expression analysis of chromosome 7 identified novel oncogene loci in non-small cell lung cancer. *Genome* 51: 1032-1039.
52. Aviel-Ronen S, Coe BP, Lau SK, da Cunha Santos G, Zhu CQ, et al. (2008) Genomic markers for malignant progression in pulmonary adenocarcinoma with bronchioloalveolar features. *Proc Natl Acad Sci U S A* 105: 10155-10160.
53. Dressman HK, Hans C, Bild A, Olson JA, Rosen E, et al. (2006) Gene expression profiles of multiple breast cancer phenotypes and response to neoadjuvant chemotherapy. *Clin Cancer Res* 12: 819-826.
54. Wang W, Nahta R, Huper G, Marks JR (2004) TAFII70 isoform-specific growth suppression correlates with its ability to complex with the GADD45a protein. *Mol Cancer Res* 2: 442-452.

Figure Legends

Figure 1. The role of alternative splicing in the TAF6 δ pathway of apoptosis. A schematic model depicts the exon 2, intron 2, exon 3 region of the *taf6* gene. Use of proximal 5' splice site (SS) generates the major TAF6 α isoform that dimerizes with its normal partner TAF9 within the TFIID complex resulting in a gene expression program allowing cell growth. Selection of the distal alternative 5' SS removes 10 amino acids to generate TAF6 δ that cannot interact with TAF9 but is incorporated into a TFIID π complex that drives a pro-apoptotic gene expression and consequently cell death.

Figure 2. TAF6 is essential for human cell viability. (A) Western blot analysis of total cell extracts 64 hours prepared 64 hours post-transfection with monoclonal antibodies specific for TAF6 α . (B) Quantification of TAF6 α protein levels from panel A using the ImageJ software package (<http://rsbweb.nih.gov/ij/>). (C) The viability of HeLa cells transfected with siRNA directed against TAF6 was measured 4 days post-transfection by methylene blue staining (see materials and methods). Viability (y-axis) is expressed relative to that of cells transfected with control siRNA. (D) TAF6 and TBP mRNA levels measured in the microarray analysis. TAF6 (dark grey) and TBP (light grey) levels detected by microarray measurements from cells transfected with siRNA that target TAF6 are shown relative to their levels in cells transfected with control siRNA.

Figure 3. Distinct impact of TAF6 δ induction versus total TAF6 mRNA depletion on the transcriptome of HeLa cells. (A) Heat map showing the impact of statistically significantly ($p < 0.05$) changes in gene expression during TAF6 mRNA depletion by siRNA at 48 and 72 hours post transfection. Red indicates induction and blue repression. Genes were ordered according to fold change at 72 hours post transfection. (B) Distribution of expression profiles amongst the six possible outcomes. Genes upregulated or downregulated at the both time points are schematized with lines. Add times (C) The previously defined TAF6 δ transcriptome signature compared to the transcriptome resulting from depletion of total TAF6 mRNA. The heat map shows the gene expression during siRNA-mediated total TAF6 mRNA depletion for the 961 TAF6 δ signature genes. (D) Venn diagrams depicting genes subsets statistically significantly regulated by total

TAF6 mRNA depletion versus by TAF6 δ induction. Upper diagram contains all regulated genes, middle diagram includes induces genes (upward arrow) and the lower Venn diagram includes repressed genes (downward arrow). (E) Gene ontology analysis of statistically significantly regulated genes during total TAF6 mRNA depletion. Enriched pathways are shown with their associated p-values.

Figure 4. Development and validation of a minigene system to study the alternative splicing of TAF6. (A) A schematic diagram showing the workflow used to study *cis*-acting RNA sequences in TAF6 alternative splicing using a TAF6 minigene plasmid. The plasmid containing an uninterrupted genomic sequence from the *taf6* gene that includes portions of exons 2 and 3, as well as the natural intron 2 is depicted along with the positions of primers used to detect exogenously expressed RNA species by RT-PCR. Minigenes were transfected into HeLa and 42 hours later total RNA was isolated for use in RT-PCR with primers from flanking plasmid sequences. PCR products were quantified by analysis using an Agilent Bioanalyzer. The percentage TAF6 δ is expressed as a ratio of total spliced TAF6 mRNAs ($\delta + \alpha$). (B) Validation of the minigene system via mutagenesis. The proximal (P) 5' splice site (SS) and distal (D) 5' SS are illustrated. Mutations that knock-out (ko) SS or strengthen by creating consensus (cons) SS and their impact on the percentage of TAF6 δ produced (x-axis), are indicated.

Figure 5. Scanning mutagenesis of constitutive exon 2 and exon 3. (A) Scanning mutations (black rectangles) in the TAF6 minigene were generated by PCR before transfection into HeLa. RNA was isolated and splice products were analysed by RT-PCR

as in Figure 4. The percentage of exogenous TAF6 δ mRNA produced by a given mutated construct are graphically shown (x-axis). (B) As in panel A, except that mutations were in exon 3.

Figure 6. Targeted mutagenesis of intron 2 of the *taf6* gene. (A) Mutations of intron motifs are indicated with arrows in the TAF6 minigene constructs. HeLa cell transfection and splice product analysis was carried out as in Figure 4. The percentage of exogenous TAF6 δ mRNA is graphically shown (x-axis). (B) As in panel A except mutations were focused on poly G motifs found within intron 2.

Figure 7. RNA secondary structure at the proximal 5' splice site can force the selection of TAF6 δ . (A) The sequence and names of mutations within alternative exon 2 (2a) of the TAF6 minigene constructs are indicated at the left. HeLa cell transfection and splice product analysis was carried out as in Figure 4. The percentage of exogenous TAF6 δ mRNA is graphically shown (x-axis). (B) As in panel A except that mutations (red nucleotides) are shown to the right in hypothetical RNA secondary structures generated using the M-Fold algorithm. The proximal 5' splice site (SS) is indicated with red boxes. (C) As in B with further mutations. (D) As in B with further mutations.

Figure 8. Evidence for an exonic splicing silencer in alternative exon 2. (A) A hypothetical RNA structure generated by M-fold is illustrated along with the position of the proximal TAF6 5' splice site (red box). Selected mutation are indicated with arrows (red text) (B) The names of mutations within alternative exon 2 (2a) of the TAF6

minigene constructs are indicated at the right. HeLa cell transfection and splice product analysis was carried out as in Figure 4. The percentage of exogenous TAF6 δ mRNA is graphically shown (x-axis). Mutations (red nucleotides) are shown to the right in hypothetical RNA secondary structures generated using the M-Fold algorithm. (C) As in panel A except that mutations (red nucleotides) are shown to the left.

Figure 9. A hypothetical model for TAF6 δ alternative splicing. (A) The TAF6 minigene construct is schematically show with putative exonic splicing enhancer (ESE), exonic splicing silencer (ESS) and intronic splicing enhancer (ISE) motifs indicated with boxes. Enhancer or silencer definitions are given with respect to the pro-apoptotic TAF6 δ isoform but likely act simultaneously to repress one 5' SS while enhancing the other because of the small distance (30 nucleotides) between them. (B) As in panel A with mutations were focused on poly G motifs found within intron 2.

Figure S1. Morphology of cell death in response to siRNA-mediated depletion of TAF6. HeLa cells were treated for 4 days with control siRNA (panel 1), or siRNA that target the TAF6 mRNA (panel 2). Cells were photographed with a phase-contrast microscope. Arrows indicate dying cells that lack classical apoptotic membrane blebbing.

Figure S2. Mapping of the branchpoint in intron 2 of the TAF6 minigene . (A) A TAF6 minigene construct bearing a point mutation in a putative branchpoint (adenine 156 to guanosine) was transfected into HeLa cells for splicing analysis as in Figure 4. The pre-mRNA and well as the spliced products are indicated with arrows. (B) Exogenously

expressed TAF6 pre-mRNA and spliced products were quantified as in Figure 4 and the percentage of exogenous TAF6 minigene splicing is graphically shown (y-axis).

Figure S3. Schematic representation of mutations used in this study. (A) The TAF6 minigene construct is shown schematically. (B) A nucleotide resolution list of mutations (altered nucleotides shown in red) from different regions of the minigene are illustrated. Black text corresponds to wild type sequences and the exons 2 and 3 (white boxes), alternative exon 2a (grey box) and intron 2 (black line) are indicated above the sequences.

Figure 1

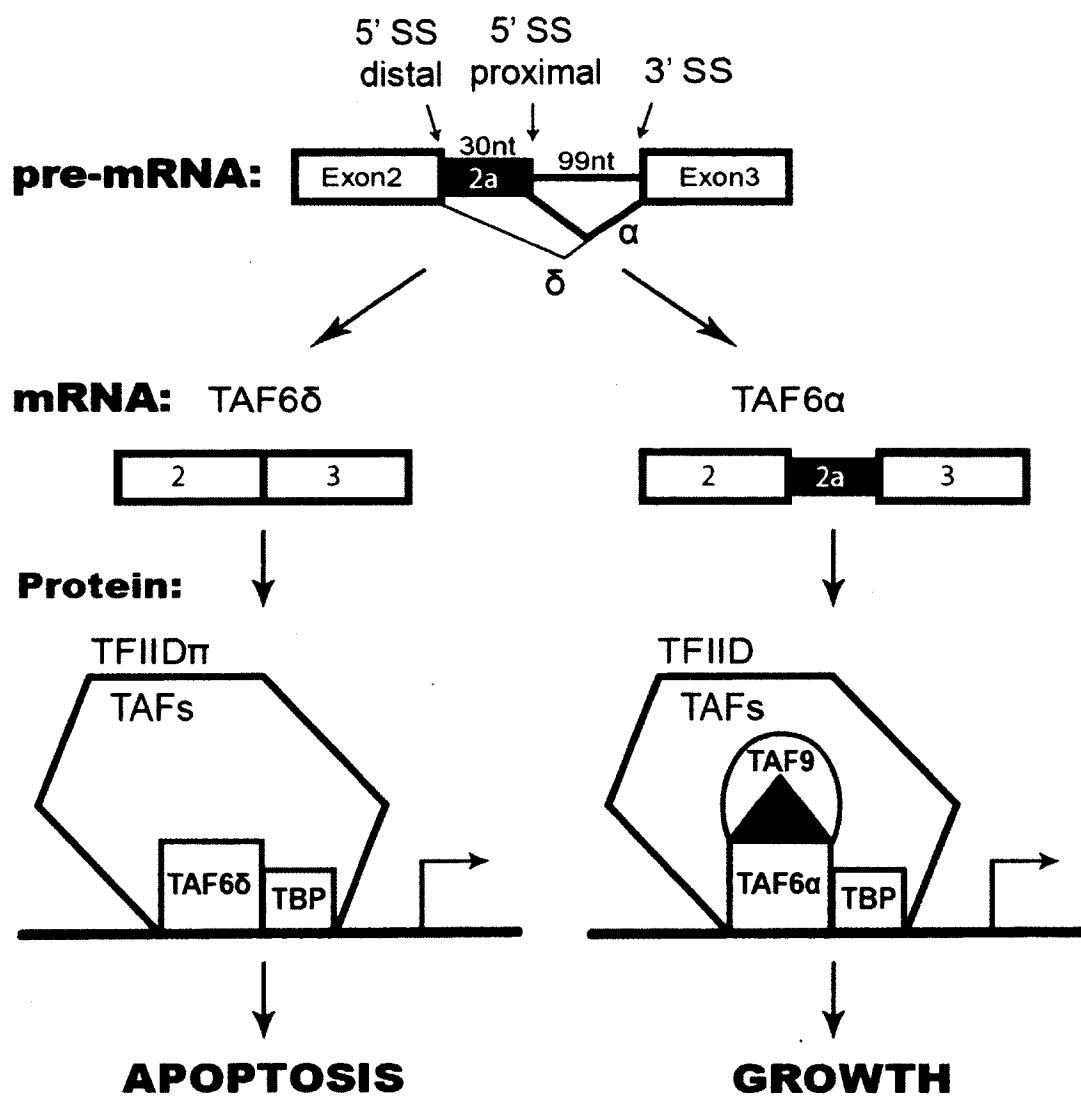


Figure 2

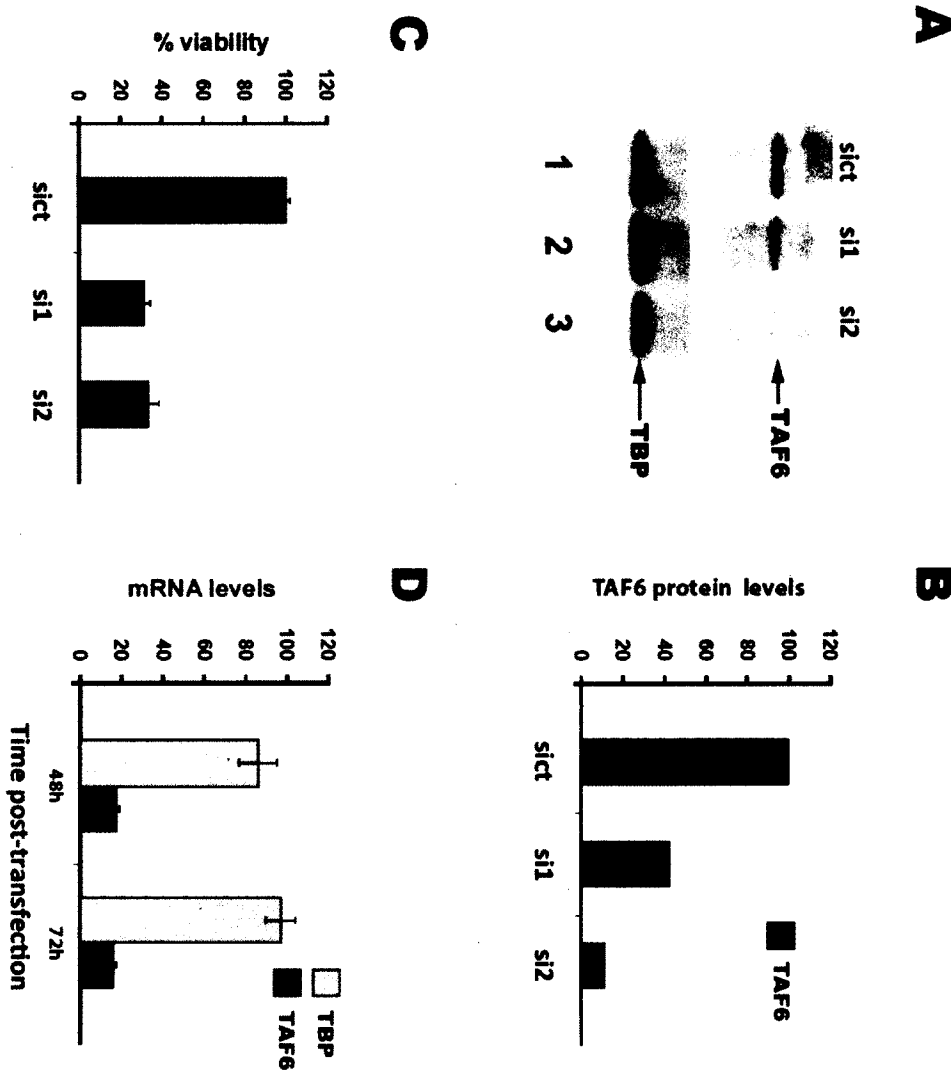


Figure 3

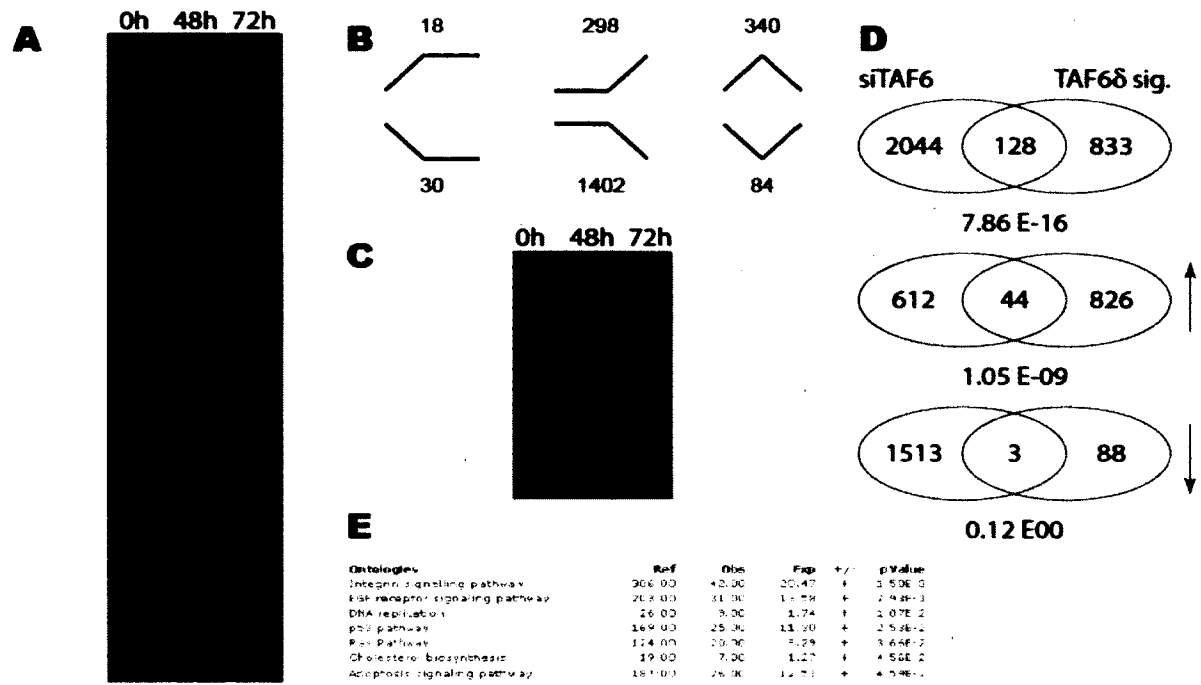


Figure 4

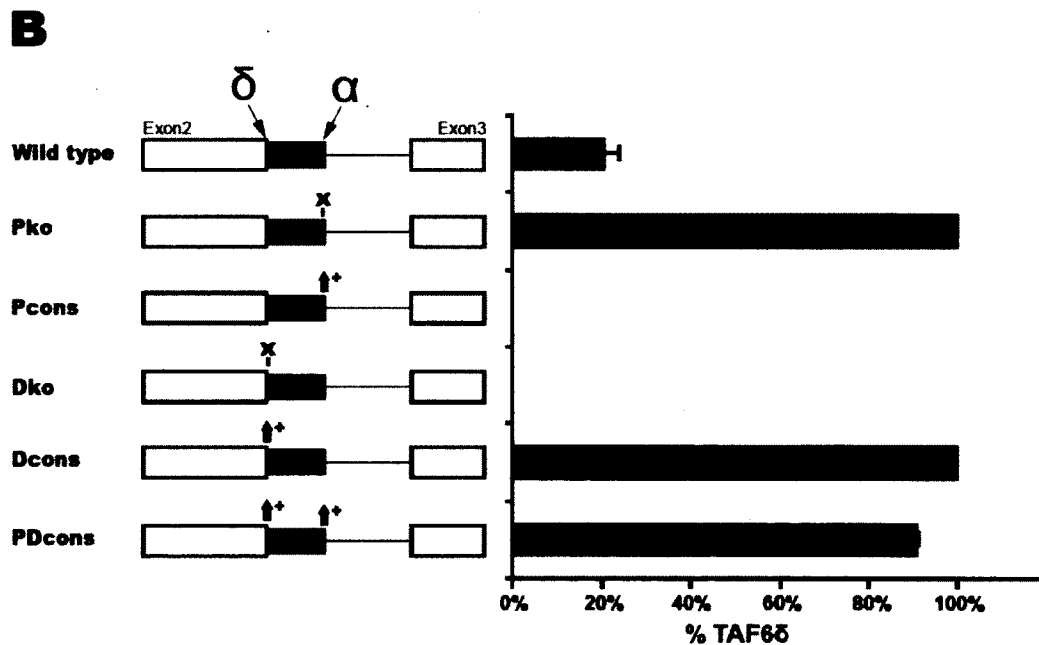
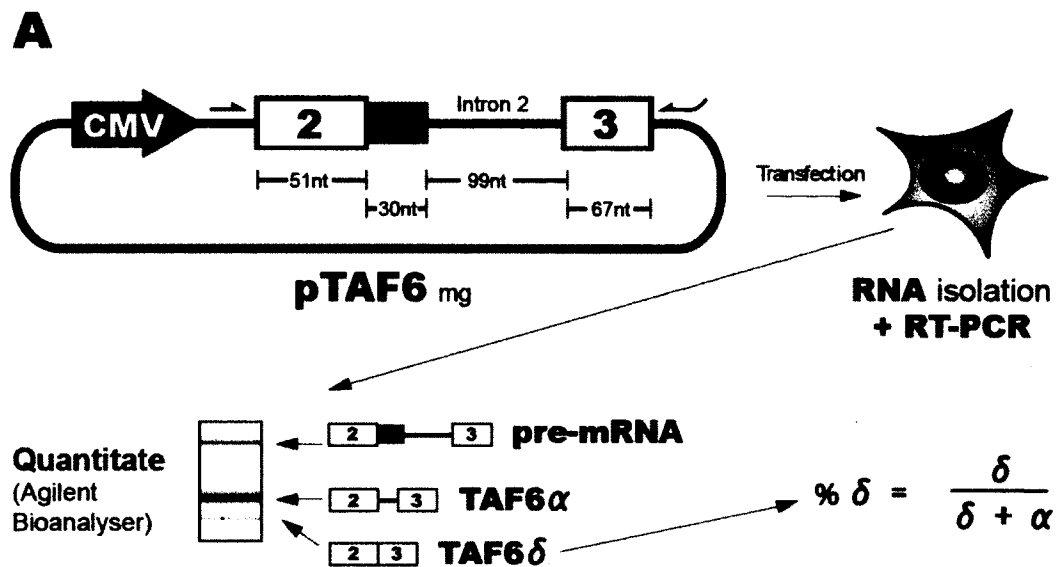
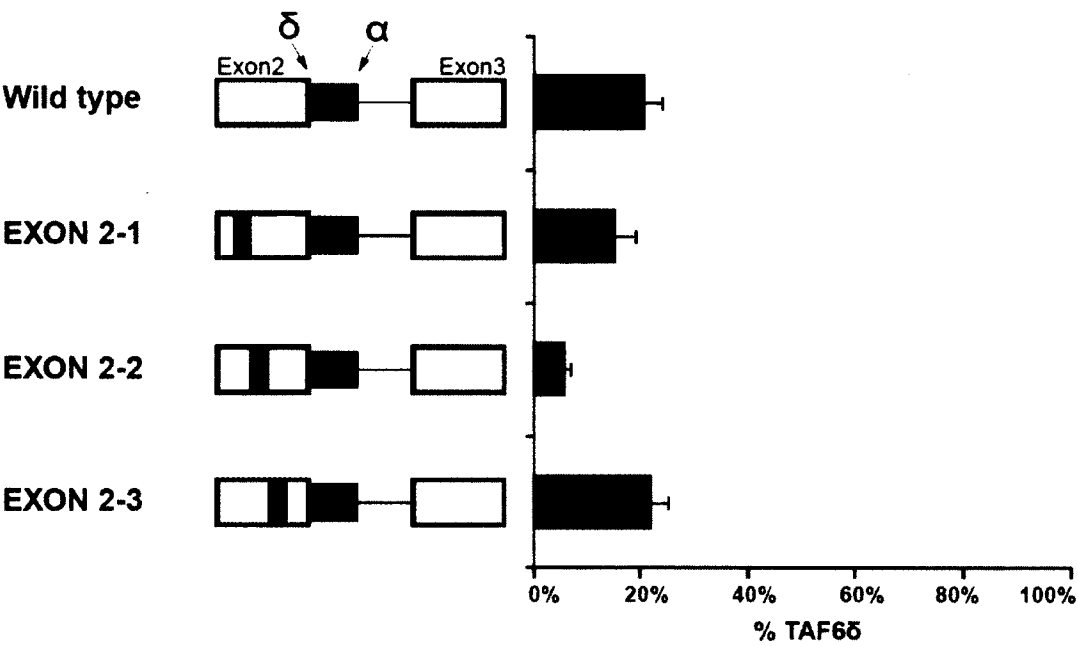


Figure 5

A



B

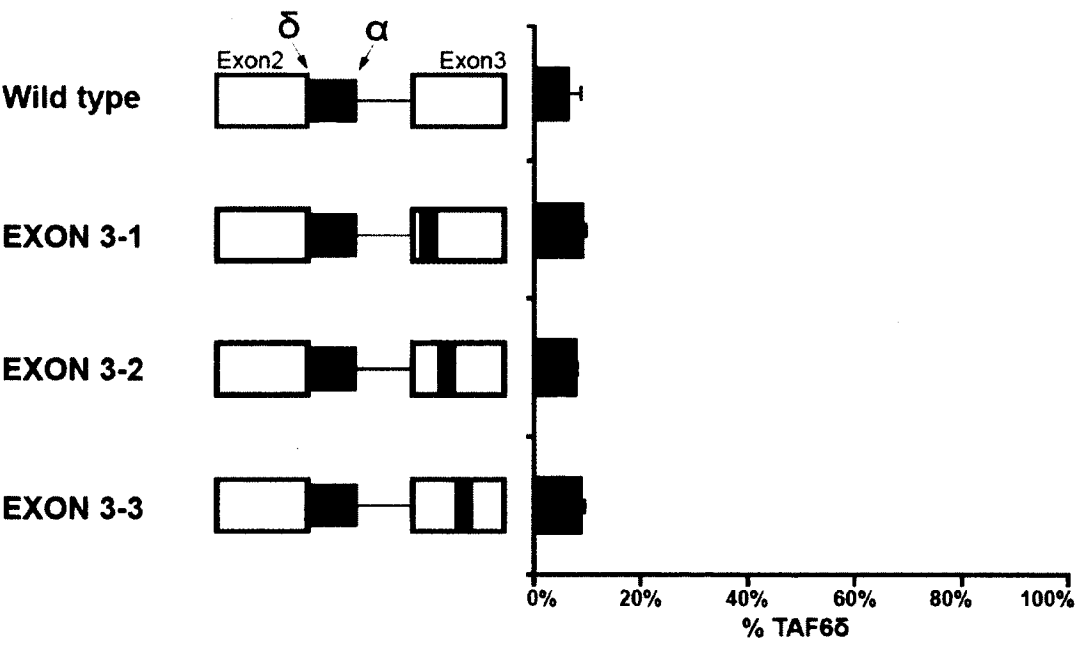


Figure 6

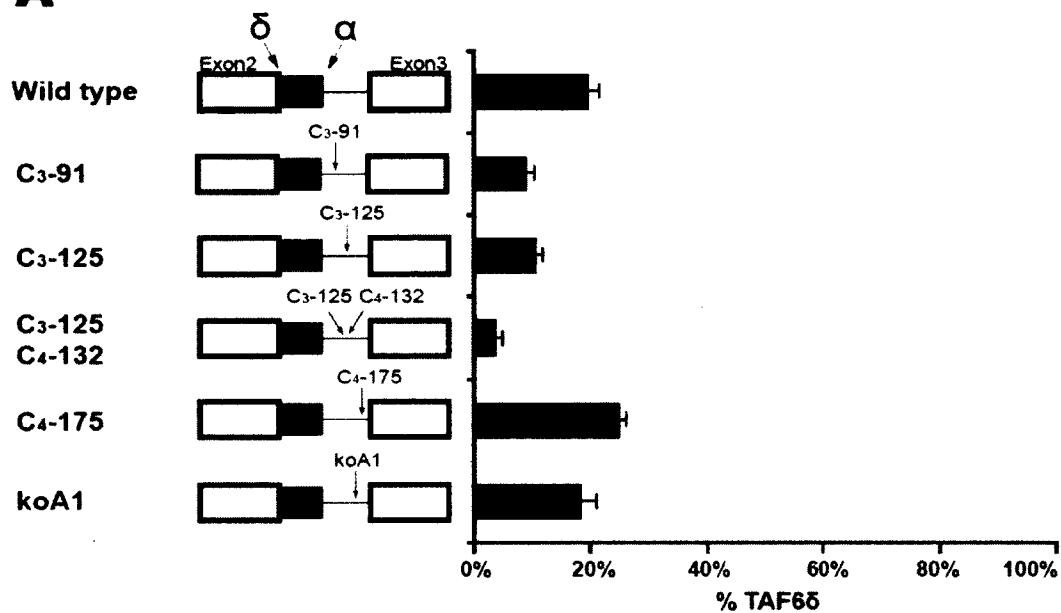
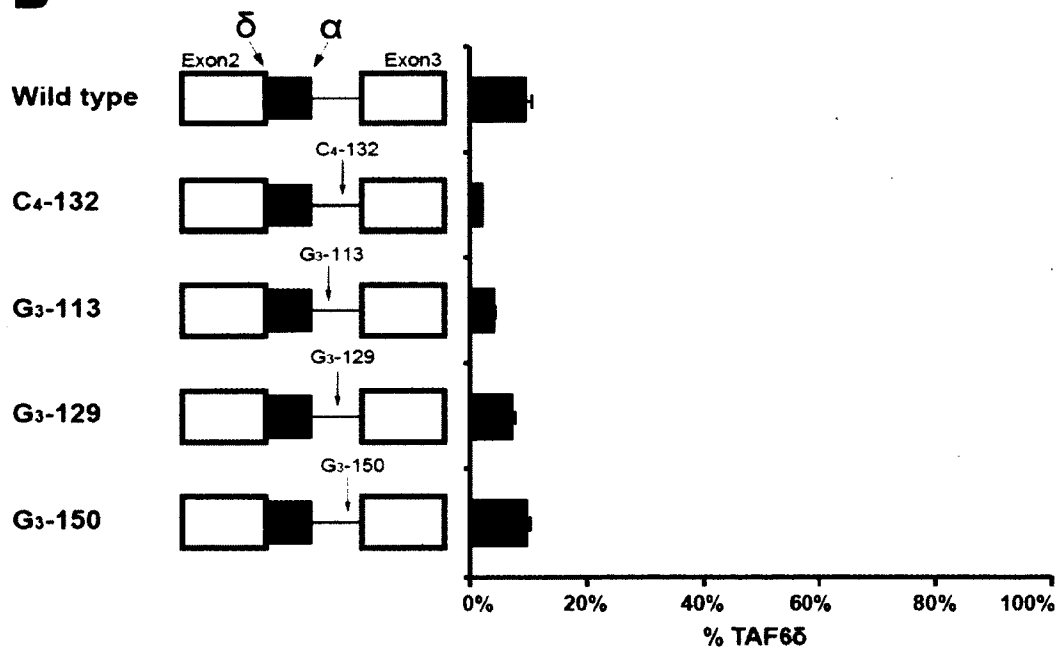
A**B**

Figure 7

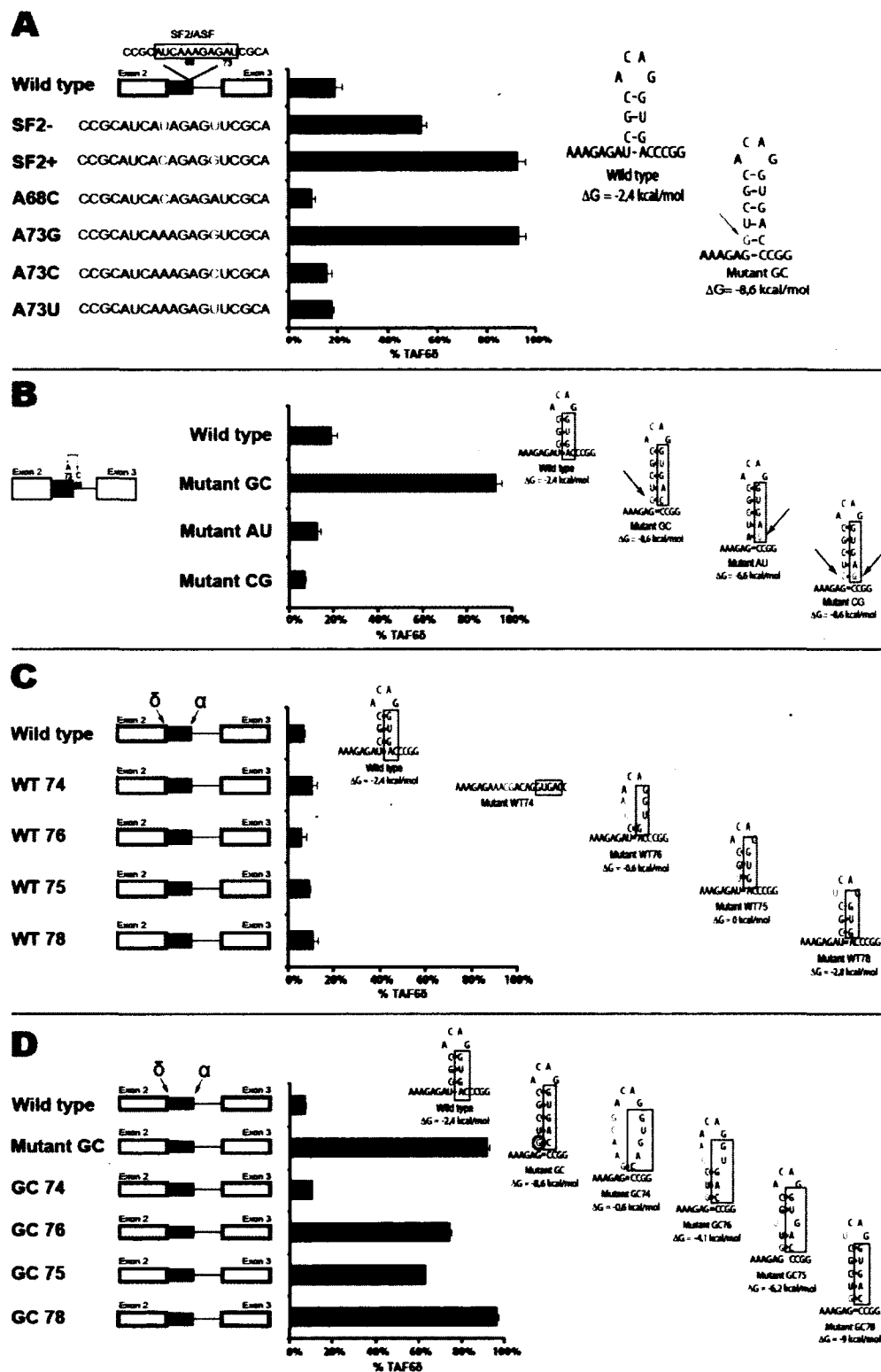


Figure 8

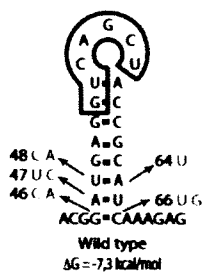
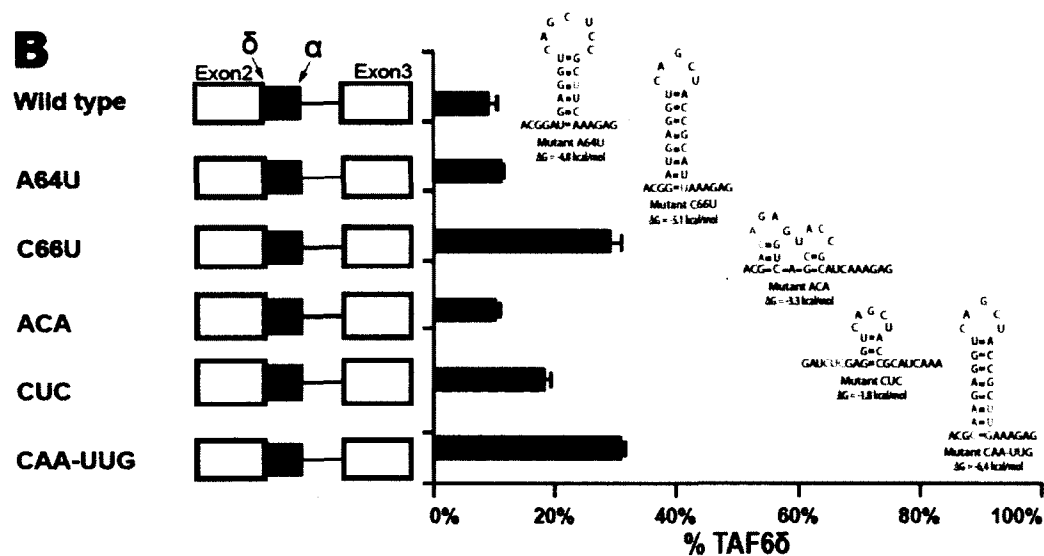
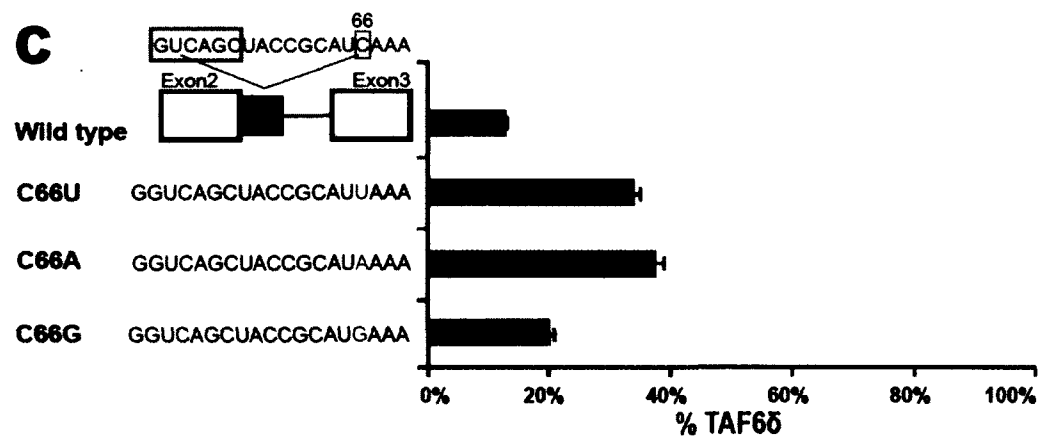
A**B****C**

Figure 9

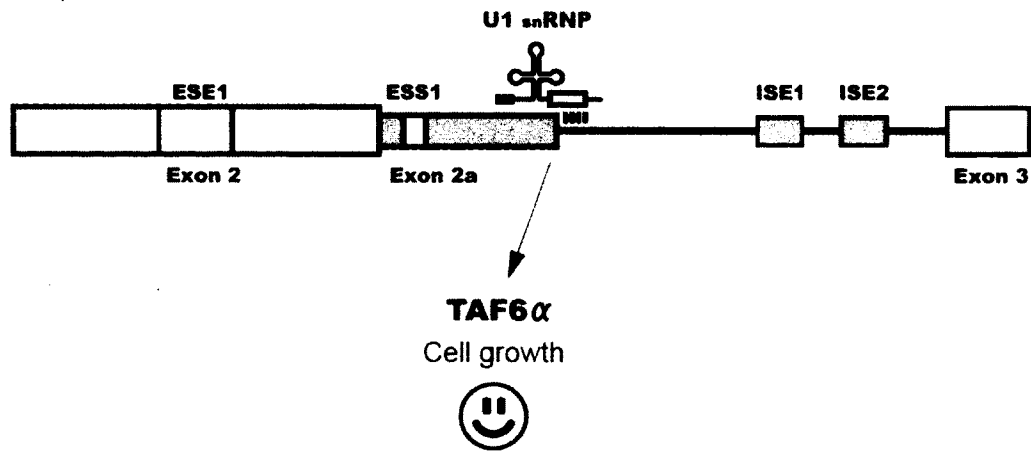
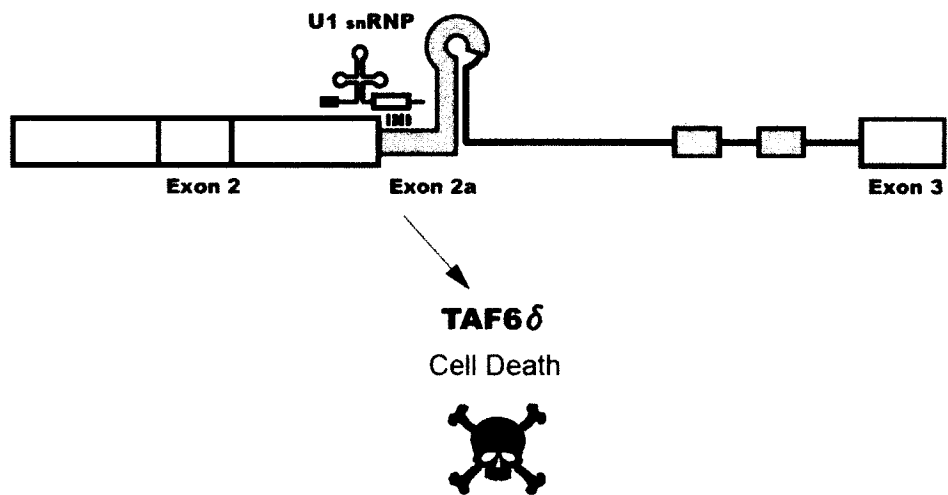
A**B**

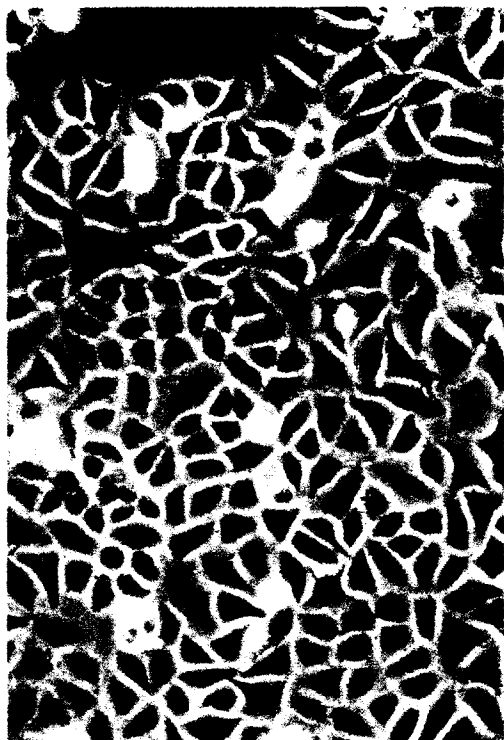
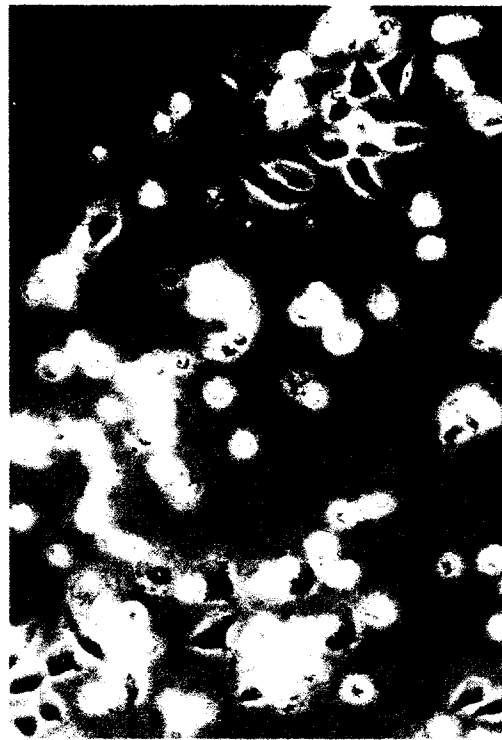
Figure S1**si Control****1****si TAF6-1****2**

Figure S2

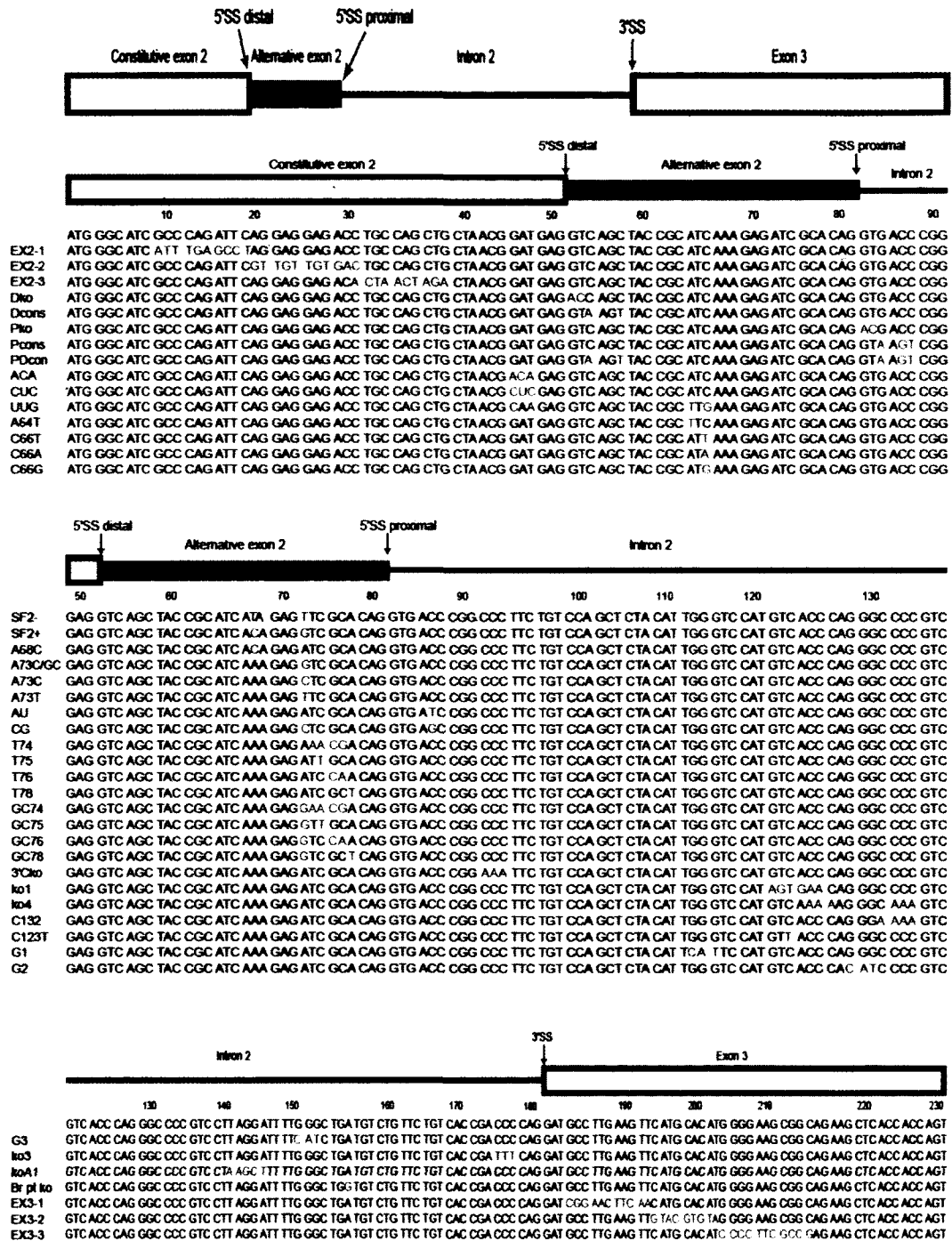
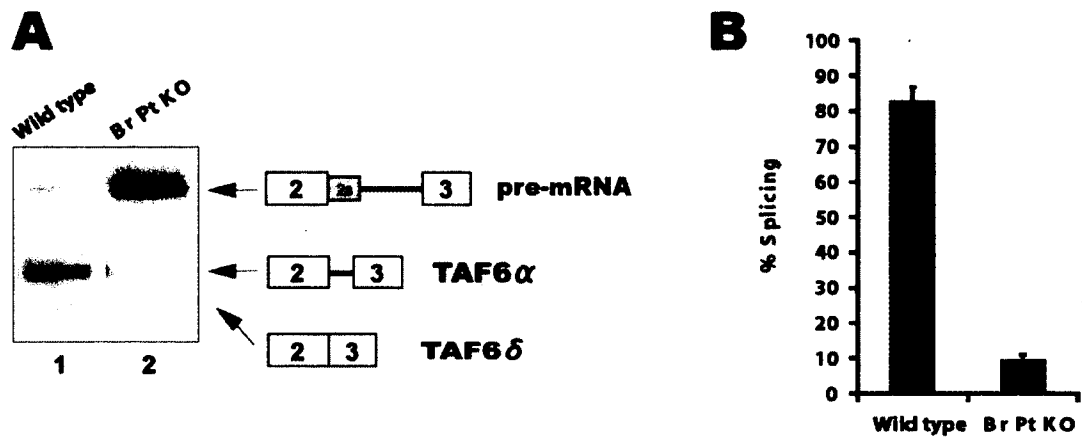


Figure S3



CHAPITRE 2

A. Les facteurs hnRNP K et hnRNP H interviennent-ils dans la régulation de l'épissage alternatif de TAF6 δ ?

Pour compléter notre étude sur le mécanisme d'épissage alternatif contrôlant l'expression de TAF6 δ versus TAF6 α , nous nous devons de connaître les facteurs qui se lient sur les éléments cis identifiés dans la première partie de mon projet.

En collaboration avec le Laboratoire de Génomique Fonctionnelle de l'Université de Sherbrooke (LGFUS), nous avons fait un criblage d'inhibition d'un certain nombre de protéines liant l'ARN. Les ARNs à interférence ont été transfectés dans sept lignées cellulaires (Hela ws, BJ, PC3, MCF7, MDA-MB-231, Skov3, Ovar3). La protéine hnRNPK a été un candidat car son inhibition augmentait l'expression de TAF6 α . D'après les résultats d'immunoprécipitation, le facteur d'épissage hnRNPK interagirait avec le pré-messager. Actuellement, nous ne savons pas si cette interaction est directe ou indirecte.

Après avoir inhibé des facteurs qui lient l'ARN, hnRNP K et hnRNP H ont été plus particulièrement surexprimées dans les cellules HeLa ws. Ces protéines sont connues pour avoir une forte affinité pour des séquences introniques polyC et polyG respectivement. L'intron 2 de TAF6 possède des séquences riches en cytidine et guanosine. Ces facteurs inhibent le plus souvent l'épissage. Notre hypothèse serait qu'une plus grande quantité de hnRNP dans les cellules HeLa ws diminuerait la quantité d'ARNm TAF6 α en comparaison à notre contrôle négatif où il n'y a pas eu de surexpression de la protéine hnRNP. L'équipe du Dr. Motorin en surexprimant la protéine hnRNP K dans les cellules a démontré que l'épissage de l'ARN pré-messager HIV-1 était inhibé (Marchand *et al.*, 2011). Les plasmides pXJ41-hnRNPK (Marchand *et al.*, 2011) et pGFP-CMV-topaz-hnRNPH (Fisette *et al.*, 2010) ont été transfectés

transitoirement dans les cellules HeLa ws à des quantités croissantes (100, 500, 750, 1000 ng) pendant 24 heures.

Des immunobuvardages de type Western ont permis de constater la surexpression de hnRNPK et hnRNPH avec des anticorps dirigés contre hnRNPK et la protéine GFP qui est fusionnée à la séquence de hnRNPH (*Figure 15*). En comparaison à notre contrôle négatif de non transfection des vecteurs d'expression des hnRNPs, une augmentation de la quantité des deux facteurs d'épissage est bien observable dans les cellules pour les quatre transfections.

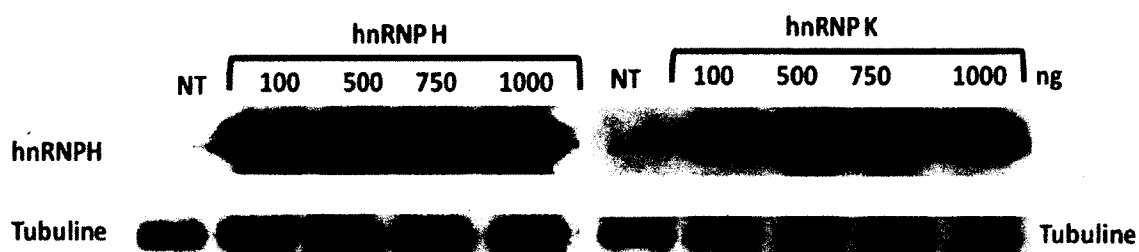


Figure 15 : La surexpression de hnRNP K et hnRNP H dans les cellules Hela ws. Immunobuvardage de type western avec des anticorps dirigés contre hnRNP H, hnRNP K et la tubuline.

Pour savoir si la surexpression des protéines a eu un impact sur l'épissage alternatif de TAF6, la présence des ARNm TAF6 α (168 nt) et TAF6 δ (138 nt) a été analysé par gel de polyacrylamide. Pour cela, vingt quatre heures après la transfection, les ARNs ont été extraits des cellules. Les ARNs ont été transcrits à l'inverse en ADNc qui par la suite ont été amplifiés par PCR avec des amorces spécifiques aux ARNs de TAF6 endogène (T61B et T62E). Enfin les ADNc amplifiés ont été déposé sur un gel de polyacrylamide qui permet par la migration de séparer les ADNc en fonction de leur taille (*Figure 16*).

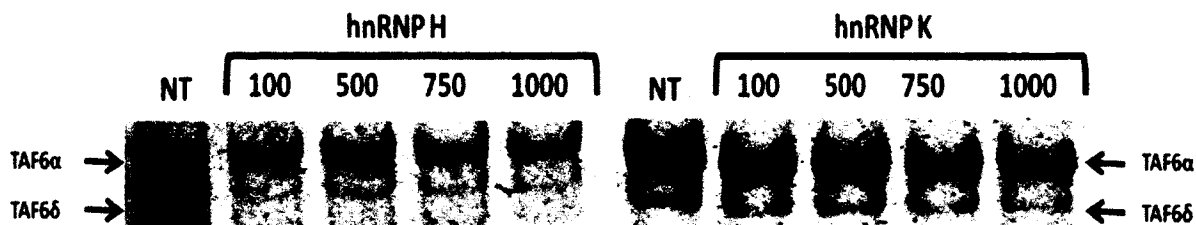


Figure 16 : Le profil d'épissage alternatif de TAF6 en fonction des quantités de hnRNP H et hnRNP K. Gel 15% non-dénaturant de RT-PCRs des ARNs messagers de TAF6 endogène dans les cellules HeLa ws transfectées par des quantités croissantes de hnRNP H et hnRNP K.

La surexpression des hnRNPs ne changent pas le profil d'épissage de TAF6 en comparaison avec le témoin négatif de non tranfection (*Figure 16*). hnRNPK et hnRNPH ne participeraient pas à la régulation de l'épissage alternatif de TAF6.

B. Développement d'un nouveau vecteur d'expression : pEYFP-TAF6-PTC

Etant donné le peu d'information que nous avons obtenu avec le criblage d'échantillon d'ADNc fournis par le LGFUS et la surexpression de hnRNP K et hnRNP H, nous avons décidé de développer un gène rapporteur nous permettant de quantifier la fluorescence des cellules exprimant ce gène rapporteur en fonction de l'expression des facteurs d'épissage. Ce système a été utilisé pour identifier les facteurs d'épissage qui régulent l'épissage du pré-messager Bcl-x (Moore *et al.*, 2010). Le pré-messager Bcl-x après l'épissage alternatif de l'exon 2 peut former une isoforme longue Bcl-xL, contenant l'exon 2 alternatif et une isoforme courte Bcl-xS, dont l'exon alternatif est excisé. Moore a utilisé un minigène correspondant à la séquence génomique du gène Bcl-x qui a été modifié par l'introduction d'un codon stop prématuré dans l'exon 2 alternatif et fusionné à la séquence d'un gène rapporteur, Venus. Les cellules qui expriment la protéine Venus émettent une fluorescence de couleur jaune en microscopie. Ce gène rapporteur fusionné permet de distinguer l'expression de Bcl-xL et Bcl-xS. En effet, lorsque Bcl-xL est le

variant produit, la machinerie de traduction se heurte au codon stop prématuré et la traduction est avortée. Les cellules n'émettent aucune fluorescence. Par contre, étant que Bcl-xS ne possède pas l'exon alternatif possédant le codon stop, il y a donc traduction de l'ARNm Bcl-xS et les cellules expriment la protéine Bcl-xS fusionnée, et émettent de la fluorescence jaune. Ce minigène rapporteur est un outil moléculaire qui est très sensible et très indicatif à l'activité des protéines d'épissage.

La présence d'un codon stop prématuré cause la synthèse d'une protéine tronquée. Ces codons stop prématurés sont reconnus par le mécanisme de surveillance de l'ARNm, le «Nonsense-Mediated mRNA Decay» (NMD) (Baserga et Benz, 1988). Le NMD protège la cellule contre les effets néfastes des protéines tronquées. Ce mécanisme dégrade les ARNm portant un codon nonsense prématuré qui crée un décalage de lecture. Le codon stop prématuré est différencié du « vrai » codon stop car il est situé à plus de 50 nucléotides en amont de la dernière jonction exon-exon. Par contre, si le codon de terminaison est situé à moins de 50 nucléotides de la dernière jonction ou dans le dernier exon, l'ARNm ne sera pas dégradé et il y aura arrêt de la traduction (Zhang *et al.*, 1998). Lors de l'épissage du pré-ARNm, les CSPs sont reconnus par le complexe protéique, EJC (pour « exon junction complex ») qui se lie 20 à 24 nucléotides en amont des jonctions exon-exon. L'EJC serait le lien entre l'épissage, l'exportation et la reconnaissance du CSP (Le Hir *et al.*, 2003).

Dans notre cas, il s'avère que l'ARN pré-messager TAF6 subit le même mode d'épissage que Bcl-x. Nous avons donc décidé de travailler avec le même système que Moore *et al.* afin d'identifier les facteurs d'épissage qui participent à l'épissage alternatif de TAF6 δ . Pour cela j'ai créé un nouveau minigène comprenant la séquence génomique de TAF6 avec un codon stop introduit par mutagenèse par PCR. Par la suite, ce minigène a été fusionné au gène EYFP pour avoir la fonction de gène rapporteur. Le principe de notre minigène rapporteur est que lorsque TAF6 α est exprimé, la traduction sera avortée, car il y a un codon stop dans l'exon2 alternatif. En revanche, les cellules qui expriment TAF6 δ émettront de la fluorescence car TAF6 δ ne possèdent pas l'exon 2 alternatif contenant le codon stop (*Figure 17*).

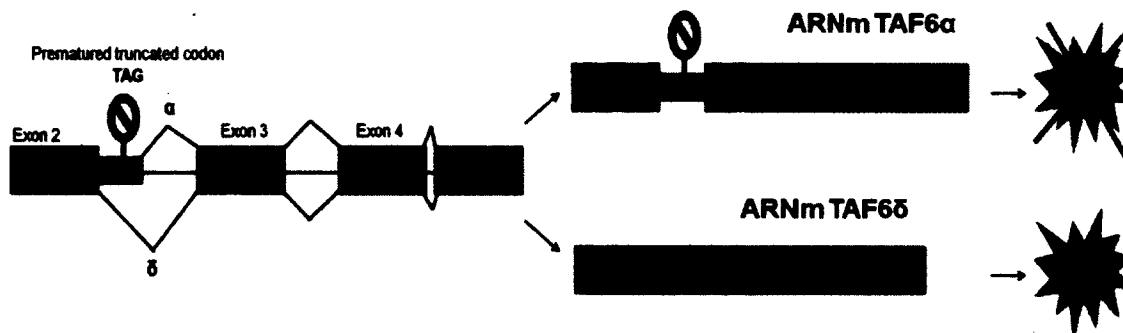


Figure 17 : Le nouveau gène rapporteur EYFP-TAF6-PTC.

Ce nouveau vecteur d'expression comprend les exons 2 à 4 de TAF6 qui sont intercalés par les exons 2 à 4 et la séquence d'EYFP. La traduction de l'ARNm TAF6 α ne permet pas de produire une protéine à cause de la présence d'un codon stop prématuré. Les cellules ne fluorescent pas. Par contre la traduction de l'ARNm TAF6 δ crée une protéine TAF6 fusionnée à EYFP qui émet de la fluorescence dans les cellules Hela ws.

Dans un premier temps, j'ai effectué des mutations dans le minigène qui nous a permis de caractériser les éléments cis participant à la régulation de l'épissage. L'introduction d'un codon stop (UAG, UGA et UAA) dans l'exon 2 alternatif ne doit pas changer le patron d'épissage. Les mutations ont été introduites par mutagenèse par PCR avec des amorces contenant les mutations (Tableau 3). Quatre mutations ont été faites dans l'exon2 alternatif (Figure 18).

	5'SS distal		Exon 2 alternatif		5'SS proximal	Intron2
	50		60	70	80	
	↓				↓	
	GAG	GTC AGC TAC CGC ATC AAA GAG ATC GCA CAG GTG ACC				
Stop 1	GAG	GTC AGC TAC CGC ATC <u>TAA</u> GAG ATC GCA CAG GTG ACC				
Stop 2	GAG	GTC AGC TAC CGC ATC <u>AAA</u> <u>TAG</u> ATC GCA CAG GTG ACC				
Stop 3	GAG	GTC AGC <u>TAG</u> CGC ATC AAA <u>GAG</u> ATC GCA CAG GTG ACC				
Stop 4	GAG	GTC AGC <u>TAC</u> <u>TGA</u> ATC AAA GAG ATC GCA CAG GTG ACC				

Figure 18 : Les différentes mutations pour introduire un codon stop dans l'exon2 alternatif de TAF6. Les codons mutés en codon stop prématurés sont soulignés. Les mutations introduites dans la séquence sont en rouge.

Le premier codon stop (UAA) a été créé au 23^{ème} codon de TAF6. L'adénosine (67^{ème} nucléotide) a été mutée en thymidine. Le deuxième codon stop (UAG) a été introduit 24^{ème} codon par la mutation de la guanosine (70^{ème} nucléotide) en thymidine. La mutation de la cytosine (60^{ème} nucléotide) en guanosine a permis d'introduire le troisième codon stop (UAG) au 20^{ème} codon. Le dernier codon stop (UGA) a été au 21^{ème} codon stop par la mutation des cytosines (61^{ème} et 63^{ème} nucléotide) en thymidine et en adénosine. Les quatre minigènes ont été transfectés dans les cellules HeLa ws transitoirement. Les ARN extraits des cellules transfectées ont été transcrits à l'inverse et amplifiés par PCR. Les ADNc ont été analysés par électrophorèse de capillarité (Agilent) afin de quantifier les isoformes d'épissage.

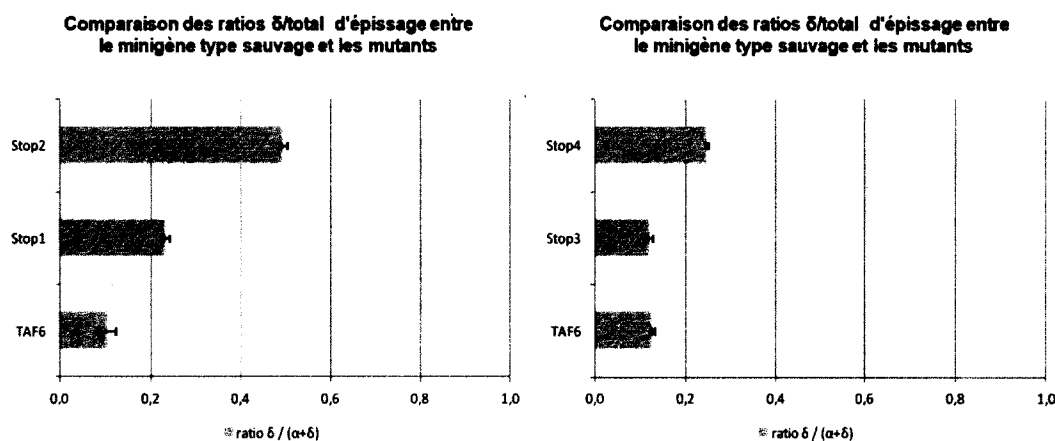


Figure 19 : Les mutations dans l'exon 2 alternatif pour introduire un codon stop. Les ratios TAF6 δ /total d'épissage (TAF6 α + TAF6 δ) ont été obtenus par le logiciel du BioAnalyser 2100 d'Agilent

Après l'obtention des données de molarité des variants d'épissage, j'ai calculé la quantité de TAF6 δ par rapport au total des produits d'épissage ($\delta/(\alpha+\delta)$). Le pourcentage de TAF6 δ obtenu après transfection des minigènes mutés a été comparé au pourcentage de TAF6 δ obtenu après une transfection du minigène de type sauvage (non muté). Les mutants stop1, stop2 et stop4 ont une quantité plus importante de TAF6 δ indiquant que les mutations aux 61^{ème}, 63^{ème}, 67^{ème}, 70^{ème} nucléotides ont perturbé la régulation de l'épissage de TAF6. Particulièrement, le mutant stop2 a environ 4 fois plus de TAF6 δ .

Nous avons démontré dans le chapitre précédent que la zone 63-70 serait un élément cis favorisant l'épissage de TAF6 α . En revanche le mutant stop3 ne modifie pas l'épissage de TAF6. Nous n'observons aucun changement de quantité que ce soit pour TAF6 δ ou TAF6 α (Figure 19). Stop 3 est donc un bon candidat pour construire notre gène rapporteur. Nous avons décidé d'effectuer la même mutation dans le plasmide pXJ42-TAF6m qui m'a servi à cloner TAF6-PTC dans le plasmide pEYFP-N1. pXJ42-TAF6m contient les exons 2, 3, 4 et les introns 2 et 3.

pEYFP-TAF6-PTC a été transfecté dans les cellules HeLa ws transitoirement. 20 h après la transfection, les cellules ont été récoltées pour observer la fluorescence émise par l'expression de EYFP-TAF6-PTC en cytométrie de flux et la production de la protéine par immunobuvardage de type western avec les anticorps appropriés. Divers contrôles m'ont permis d'analyser la transfection pEYFP-TAF6-PTC. Les contrôles sont 1) les cellules non transfectées, 2) la transfection de pEYFP-N1, le plasmide où a été cloné TAF6-PTC (Figure 20 C), 3) la transfection de pEYFP-TAF6, TAF6 ne contient pas de codon stop prématuré dans l'exon 2 alternatif (Figure 20 B), 4) la co-transfection de pEYFP-TAF6-PTC (Figure 20 A) et l'oligonucléotide antisens TAF6 AS1 qui en se liant à l'exon 2 alternatif, force l'épissage alternatif en faveur de TAF6 δ et 5) la co-transfection de pEYFP-TAF6-PTC et l'oligonucléotide antisens contrôle.

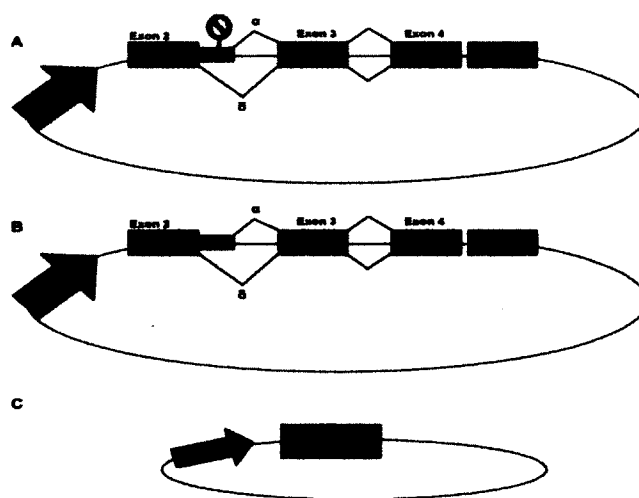


Figure 20 : Les plasmides dérivés de pEYFP-N1. A) pEYFP-TAF6-PTC. B) pEYFP-TAF6. C) pEYFP-N1.

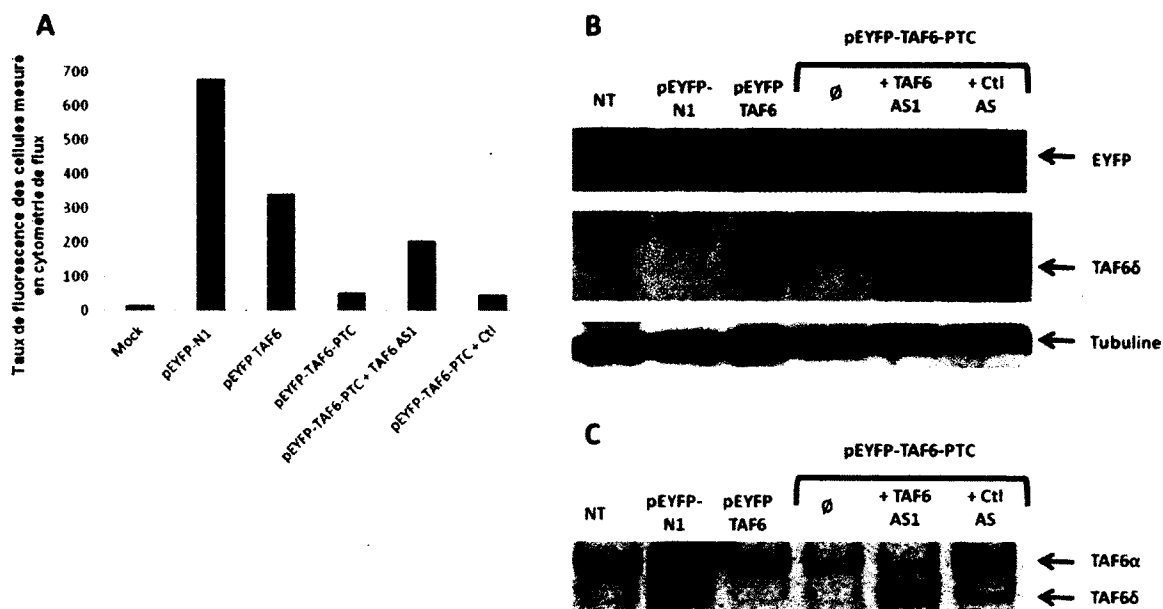


Figure 21 : L'expression de pEYFP-TAF6-PTC dans les cellules HeLa ws.

A) Fluorescence émise par les cellules HeLa ws transfectées par les différents plasmides. B) immunobuvardage de type western des différentes transfections des cellules HeLa ws avec des anticorps dirigés contre EYFP, TAF6δ et la tubuline. C) Gel 15% non-dénaturant de RT-PCRs des ARNs messagers de TAF6 endogène et exogène dans les cellules HeLa ws transfectées.

Pour analyser la fluorescence émise par les cellules, j'ai pris en compte les moyennes de la fluorescence. Les résultats sont sous forme de diagramme (Figure 21 A). Notre témoin de transfection Mock ne présente aucune fluorescence puisqu'il n'y a pas eu de transfection. Le deuxième contrôle pEYFP-N1 présente un taux de fluorescence maximal en comparaison avec pEYFP-TAF6 et pEYFP-TAF6-PTC. L'insertion de TAF6 dans le plasmide pEYFP-N1 diminue le taux de fluorescence des cellules. Le taux de fluorescence des cellules transfectées par pEYFP-TAF6-PTC est moins important que celles transfectées par pEYFP-TAF6. La présence du codon stop prématuré a donc diminué l'expression de TAF6 fusionné à EYFP. Notre contrôle positif est la co-transfection de pEYFP-TAF6-PTC et TAF6 AS1. La transfection de l'ARN antisens TAF6 AS1 permet la formation de plus d'ARNm mature TAF6δ et par conséquent augmente l'expression de TAF6δ. En comparaison avec la transfection de pEYFP-TAF6-

PTC seul, on constate que l'ajout de TAF6 AS1 permet l'augmentation de la fluorescence émise par les cellules transfectées. Enfin notre dernier contrôle nous a permis d'observer l'efficacité de la transfection de l'ARN antisens TAF6 AS1. Le taux de fluorescence est le même que lorsque pEYFP-TAF6-PTC a été transfecté seul. Les résultats de fluorescence démontrent que le système fonctionne.

Pour appuyer nos résultats de fluorescence, j'ai effectué un immunobuvardage de type Western pour observer l'expression de TAF6 δ dans les différents cas de transfections. L'anticorps dirigé contre TAF6 δ m'a permis de voir son expression (Figure 21 B). J'ai constaté l'expression de TAF6 δ dans le cas de la transfection de pEYFP-TAF6. Par contre, nous voyons TAF6 δ à une très faible quantité dans le cas où les cellules ont été transfectées par pEYFP-TAF6-PTC. Lorsqu'il y a induction de TAF6 δ par le SSO, il y a plus de protéines en comparaison avec la piste pEYFP-TAF6. Nos contrôles négatifs Mock et pEYFP-N1 n'ont pas de TAF6 δ d'après l'immunobuvardage. Le contrôle est surtout un contrôle de fluorescence des cellules. Le contrôle de co-transfection de pEYFP-TAF6-PTC et de l'ARN antisens contrôle montre que l'effet d'ARN antisens TAF6 AS1 est bien spécifique à l'induction de TAF6 δ .

Enfin, l'analyse des différentes transfections a été complétée par l'observation du patron d'épissage au niveau des ARNm (Figure 21 C). Les ARNs extraits des échantillons ont été transcrits à l'inverse et amplifiés avec des amorces spécifiques au plasmide pEYFP-TAF6-PTC. Les ADNc qui ont migré sur gel de polyacrylamide m'ont permis de constater la formation de TAF6 α versus TAF6 δ après épissage alternatif et particulièrement l'efficacité de l'oligonucléotide antisens TAF6 AS1. Dans les contrôles des cellules non transfectées, transfectées par pEYFP-N1 et pEYFP-TAF6, TAF6 α est l'isoforme majoritaire. La piste correspondant à la transfection de TAF6 AS1 montre qu'il y a plus d'ARNm TAF6 δ . Nous pouvons dire que TAF6 AS1 a été efficace et que notre contrôle positif est bien fonctionnel en comparaison à la transfection de pEYFP-TAF6-PTC

	et	l'ARN
	antisens	contrôle.

DISCUSSION ET CONCLUSION

Un des objectifs de l'équipe du professeur Brendan BELL, est de comprendre le mécanisme d'activation de l'apoptose par l'isoforme pro-apoptotique TAF6 δ , et notamment de définir le mécanisme d'épissage alternatif de TAF6 δ . Pour cela, il est important de connaître les éléments *cis*. De plus, comme l'épissage est vraisemblablement régulé par des protéines qui lient ces éléments en question, il serait important de les identifier.

Donc le premier objectif de mon projet a été de caractériser les séquences *cis* qui régulent l'épissage alternatif en faveur de TAF6 δ . Notre minigène constitué des exons 2 et 3 et l'intron 2 du gène TAF6 a permis de mener cette étude. Dans une première étape, notre outil moléculaire a été transfecté dans les cellules HeLa ws afin de quantifier la production des ARN messagers TAF6 α et TAF6 δ donc de valider notre minigène (Wilhelm *et al.* 2008). Par la suite, les sites d'épissage proximal (Pko, Pcons) et distal (Dko, Dcons) de l'exon 2 de TAF6 ont été mutés pour tester leur force relative au cours de l'épissage (*Figure 4* de l'article). Il s'avère que le site d'épissage proximal (alpha) est un site fort par rapport au site distal (delta) car dans le cas du type sauvage, l'ARNm TAF6 α est le variant d'épissage majoritaire (90%). Etant donné que les profils d'épissage des TAF6 endogène et exogène sont similaires et les deux sites d'épissage bien reconnus par la machinerie d'épissage (spliceosome), le minigène TAF6 est bien l'outil qui nous permettra d'identifier les éléments *cis* du pré-messenger. De plus, la mutation du site de branchement (Br pt ko) qui diminue considérablement le taux d'espèces épissés, renforce la validation de notre minigène pour effectuer une analyse mutationnelle (*Figure S3* du chapitre 1).

Des mutations aléatoires dans l'exon 2 et 3 nous ont permis de constater que la deuxième séquence mutée de l'exon 2 (ex2_2) influencerait l'épissage en faveur de TAF6 δ (*Figure 5* du chapitre 1). Par contre les mutations dans l'exon 3 n'ont pas changé les profils d'épissage (*Figure 5* du chapitre 1). Ce résultat était attendu car en général l'exon qui suit les exons épissés ne participe pas à la régulation de l'épissage alternatif de sites accepteurs.

Le choix de certaines mutations a été guidé par un outil de recherche de sites putatifs reconnus par des facteurs d'épissage. Le logiciel ESE Finder (<http://exon.cshl.edu/ESE/>) permet d'analyser le pré-messager pour identifier des sites potentiellement liés par les protéines SR (SF2/ASF, SC35, SRp40, SRp55) (Cartegni *et al.* 2003). ESE Finder compile les différentes données de SELEX (Selected Evolution of Ligands through EXponential enrichment), une méthode de sélection de sites de liaison à haute affinité à partir d'une bibliothèque d'ARN (Tuerk et Gold 1990). Une séquence dans l'exon 2 alternatif est susceptible d'être liée par SF2/ASF (*Figure 7A* du chapitre 1). Les mutations sur ce site n'ont pas été informatives. La mutation sur un site intronique susceptible d'être reconnu par hnRNP A1 (koA1) n'a pas perturbé l'épissage de TAF6 (*Figure 6* et *S2* du chapitre 1) (Fisette *et al.* 2010).

A l'issu de ces mutations du site hypothétiquement lié par SF2/ASF, nous avons pu constater qu'une des mutations avait créé une structure secondaire qui basculait radicalement l'épissage en faveur de TAF6δ (*Figure 7A* du chapitre 1). Des mutations complémentaires ont démontré que la mutation GC est bien particulière et propre à la formation d'une conformation tige-boucle (*Figure 7C* et *7D* du chapitre 1). Dans certains cas, la structure secondaire intervient dans la régulation de l'épissage comme pour l'épissage de l'exon 7 du gène SMN2. La tige-boucle proche du site 5' d'épissage de l'exon empêche la fixation du snRNP U1, réduisant le taux d'inclusion de cet exon (Singh *et al.* 2007). Nous savons qu'au cours de l'épissage, le pré-messager peut exister sous différentes conformations. La présence d'une tige boucle peut activer l'épissage en rapprochant les sites 5' et 3' d'épissage très éloignés par la taille des introns (Warf et Berglund 2010). Dans notre cas, la présence de la structure secondaire dans le mutant GC empêche la fixation du snRNP U1 sur le site 5' d'épissage proximal (alpha) en le séquestrant.

Le logiciel Mfold utilise des algorithmes qui donnent des prédictions de structure secondaire dans les ARNs. L'hypothèse de la présence d'une tige-boucle rapprochant l'exon 2 constitutif et alternatif nous a intéressés. Le résultat des mutations pour casser (ACA, CUC), restaurer (CAA-UUG) et tester la stabilité (A64U, C66U) de la tige-boucle

n'est pas allé dans le sens de la diminution d'ARNm TAF6 δ (*Figure 8A* du chapitre 1). Cette zone du minigène n'est pas sujette à la formation d'une structure secondaire. En revanche, la mutation C66U a augmenté de trois fois la quantité d'ARNm TAF6 δ . J'ai voulu savoir si cette mutation ponctuelle avait les mêmes effets avec d'autres nucléotides (C66A et C66G). La mutation C66A a un effet un peu plus important que C66U. La mutation a dû perturber un site de liaison d'un facteur de liaison. Le nucléotide 66 ferait partie d'une séquence *cis* qui inhiberait l'épissage de TAF6 δ (ESS) (*Figure 8B* du chapitre 1).

Les motifs polyC et polyG sont connus pour être liés par les protéines hnRNP K et H respectivement (Makeyev et Liebhaber 2002 ; Caputi et Zahler 2001). L'intron 2 de TAF6 possède plusieurs motifs (Martinez-Contreras *et al.*, 2006). Seules les mutations des motifs polyC 132 et polyG 113 ont diminué le taux de l'ARNm TAF6 δ . Ces deux motifs introniques activeraient l'épissage de TAF6 δ (ISE). Au vue de ces résultats, nous avons pu émettre un modèle de la régulation de l'épissage alternatif (*Figure 9* du chapitre 1).

A la lumière de ces résultats, nous nous posons la question, quelles sont les facteurs d'épissage liant ces éléments *cis*. Nous avons donc voulu savoir si hnRNP K et H étaient impliqués dans l'épissage. Pour cela, ces protéines ont été surexprimées et inhibées dans les cellules HeLa ws. Dans les deux cas, le profil d'épissage de TAF6 n'a pas changé. Les essais d'inhibition de ces deux protéines n'ont pas eu beaucoup d'effet sur le patron d'épissage de TAF6. Ainsi, afin d'approfondir notre étude, j'ai généré un nouvel outil moléculaire ayant les propriétés (1) d'observer les effets de la surexpression et l'inhibition des facteurs en trans sur l'épissage en faveur de TAF6 δ dans les cellules (2) la facilité de manipuler une lignée stable exprimant ce nouveau vecteur d'expression.

Dans cette deuxième partie, je me suis focalisée sur la création d'un outil moléculaire qui nous permettra de connaître quels sont les facteurs d'épissage qui influencent l'épissage alternatif de TAF6 δ . Ce gène rapporteur correspond à la séquence génomique de TAF6 (exon 2 à 4 et intron 2-3) fusionnée à la protéine fluorescente EYFP avec un codon stop

prématuré (*Figure 17*). La présence du CSP, dans l'exon 2 alternatif de TAF6, permet de distinguer les cellules dont l'épissage a produit des ARNm TAF6 α aux cellules qui possèdent des ARNm TAF6 δ . L'ARNm TAF6 α va être vite dégradé par les nucléases, car il présente un codon stop prématuré. Par contre, l'ARNm TAF6 δ va être traduit normalement et la protéine exprimée va permettre aux cellules d'émettre de la fluorescence de couleur verte. Cette fluorescence est observable au microscope optique et quantifiable par cytométrie de flux. Cette démarche a permis à Moore *et al.* de montrer que l'inhibition de SF3B1 favorisait l'épissage de la forme pro-apoptotique de Bcl-xS car les cellules expriment Bcl-xS fusionné au gène Venus qui fluoresce dans le vert (Moore *et al.* 2010).

Nos essais de transfections transitoires dans les cellules HeLa ws ont montré que notre système moléculaire est efficace puisque nous observons une différence de fluorescence entre les cellules qui expriment TAF6 α (pas ou peu de fluorescence, pEYFP-TAF6-PTC seul) et celles qui expriment TAF6 δ (au moins deux fois plus de fluorescence, pEYFP-TAF6-PTC + TAF6 AS1) (*Figure 21*). Nos résultats ont été complétés par des immunobuvardages de type western avec des anticorps dirigés contre TAF6 δ et EYFP.

Ainsi, l'une des approches futures consisterait à exprimer stablement notre construction dans les cellules HeLa ws. Cette lignée nous permettra de faire un criblage de gain de fonction (cDNA) et d'inhibition (siRNA) de protéines connues pour se lier à l'ARN dans un premier temps. L'équipe du Dr Cartens a voulu identifier les protéines qui participent à la régulation de l'épissage de FGFR2 par cette même approche. Deux minigènes exprimant les protéines fluorescentes EGFP et mRFP dépendent de l'épissage de l'exon IIIb et l'exon IIIc de FGFR2 respectivement. Un criblage fonctionnel a été entrepris sur les cellules DT3 et AT3 exprimant stablement ces deux minigènes (Newman *et al.* 2006). Bien entendu, une approche biochimique sera primordiale pour connaître leurs sites de liaison. Des études d'immunoprécipitation d'ARN ont donné des résultats préliminaires. Cette technique se doit d'être améliorée pour approfondir notre étude du mécanisme d'épissage alternatif du facteur de transcription TAF6.

L'épissage permet la biodiversité à partir d'un nombre restreint de gènes, l'existence d'un tel processus permet une régulation spécifique de l'apoptose qui mène un contrôle rapide et efficace de son activité, puisqu'à partir du même pré-ARNm, il y a la possibilité de produire soit des protéines actives ou à l'inverse des protéines inhibitrices (Schwerk et Schulze-Osthoff 2005). La régulation de l'épissage résulte d'une balance constituée d'un côté des facteurs activateurs et d'un autre côté des facteurs inhibiteurs. L'augmentation ou la diminution des concentrations de l'un de ces facteurs peut modifier le profil d'épissage. Peu d'études se sont intéressées à TAF6 δ . La compréhension du rôle de cette nouvelle voix apoptotique ainsi que le signal qui active cette voix revêt donc une importance physiologique ainsi que pharmacologique majeure car c'est une voix qui nous permet de développer de potentielles cibles thérapeutiques anti-cancéreuses.

Au total, notre étude aura permis de (1) mieux caractériser les éléments cis de l'ARN pré-messager TAF6 impliqués dans la régulation de l'épissage alternatif de TAF6 δ et de (2) générer un gène rapporteur présentant un potentiel intérêt pour le développement d'une approche moléculaire d'identification de facteur d'épissage.

REMERCIEMENTS

Je voudrais remercier le Dr. Brendan BELL de m'avoir accueilli dans son laboratoire et de m'avoir permis de mener ce projet. Ce fût pour moi un grand plaisir d'apprendre, de performer dans cet environnement scientifique. Je remercie beaucoup Emmanuelle WILHELM et Sébastien CAGNOL, plus communément Manou et Babao, pour leurs écoutes scientifiques, leurs soutiens, leurs joies de vivre. Je remercie Aurélie DELANNOY et Edith ALVARADO pour le soutien et ce partenariat qui nous a fait grandir. Je remercie aussi Alexandre MADGINSKI et Julien BOUDREAULT pour leurs futures mains et cerveaux de chercheurs avertis.

Je remercie les Dr Benoit CHABOT, Dr Sherif ABOU ELELA et Dr MOTORIN et leur équipe pour ce coup de pouce technique qui a permis de faire avancer mon projet. Je remercie également à Dr Martin BISAILLON d'avoir corrigé mon mémoire.

Je terminerai par exprimer mon amour et ma gratitude à ma famille et mes proches qui me supportent, me soutiennent, m'encouragent à me surpasser.

REFERENCES

- Aoyagi, N. and D. A. Wassarman (2001). "Developmental and transcriptional consequences of mutations in *Drosophila* TAF(II)60." *Mol Cell Biol* 21(20): 6808-19.
- Ashkenazi, A. (2008). "Targeting the extrinsic apoptosis pathway in cancer." *Cytokine Growth Factor Rev* 19(3-4): 325-31.
- Aviel-Ronen, S., B. P. Coe, et al. (2008). "Genomic markers for malignant progression in pulmonary adenocarcinoma with bronchioloalveolar features." *Proc Natl Acad Sci U S A* 105(29): 10155-60.
- Bando, M., S. Ijuin, et al. (1997). "The involvement of the histone fold motifs in the mutual interaction between human TAF(II)80 and TAF(II)22." *J Biochem* 121(3): 591-7.
- Baserga, S. J. and E. J. Benz, Jr. (1988). "Nonsense mutations in the human beta-globin gene affect mRNA metabolism." *Proc Natl Acad Sci U S A* 85(7): 2056-60.
- Bell, B., E. Scheer, et al. (2001). "Identification of hTAF(II)80 delta links apoptotic signaling pathways to transcription factor TFIID function." *Mol Cell* 8(3): 591-600.
- Bellorini, M., J. C. Dantonel, et al. (1996). "The major histocompatibility complex class II Ea promoter requires TFIID binding to an initiator sequence." *Mol Cell Biol* 16(2): 503-12.
- Berget, S. M. (1995). "Exon recognition in vertebrate splicing." *J Biol Chem* 270(6): 2411-4.
- Berget, S. M., C. Moore, et al. (1977). "Spliced segments at the 5' terminus of adenovirus 2 late mRNA." *Proc Natl Acad Sci U S A* 74(8): 3171-5.
- Birck, C., O. Poch, et al. (1998). "Human TAF(II)28 and TAF(II)18 interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family." *Cell* 94(2): 239-49.
- Black, D. L. (2003). "Mechanisms of alternative pre-messenger RNA splicing." *Annu Rev Biochem* 72: 291-33.
- Boutz, P. L., P. Stoilov, et al. (2007). "A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons." *Genes Dev* 21(13): 1636-52.
- Bradsher, J. N., K. W. Jackson, et al. (1993). "RNA polymerase II transcription factor SIII. I. Identification, purification, and properties." *J Biol Chem* 268(34): 25587-93.

- Bradsher, J. N., S. Tan, et al. (1993). "RNA polymerase II transcription factor SIII. II. Functional properties and role in RNA chain elongation." *J Biol Chem* 268(34): 25594-603.
- Breathnach, R. and P. Chambon (1981). "Organization and expression of eucaryotic split genes coding for proteins." *Annu Rev Biochem* 50: 349-83.
- Buratowski, S., S. Hahn, et al. (1989). "Five intermediate complexes in transcription initiation by RNA polymerase II." *Cell* 56(4): 549-61.
- Burke, T. W. and J. T. Kadonaga (1996). "Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters." *Genes Dev* 10(6): 711-24.
- Burke, T. W. and J. T. Kadonaga (1997). "The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila." *Genes Dev* 11(22): 3020-31.
- Burley, S. K. and R. G. Roeder (1996). "Biochemistry and structural biology of transcription factor IID (TFIID)." *Annu Rev Biochem* 65: 769-99.
- Burset, M., I. A. Seledtsov, et al. (2001). "SpliceDB: database of canonical and non-canonical mammalian splice sites." *Nucleic Acids Res* 29(1): 255-9.
- Buschmann, T., O. Potapova, et al. (2001). "Jun NH2-terminal kinase phosphorylation of p53 on Thr-81 is important for p53 stabilization and transcriptional activities in response to stress." *Mol Cell Biol* 21(8): 2743-54.
- Caceres, J. F., S. Stamm, et al. (1994). "Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors." *Science* 265(5179): 1706-9.
- Campbell, J. M., W. W. Lockwood, et al. (2008). "Integrative genomic and gene expression analysis of chromosome 7 identified novel oncogene loci in non-small cell lung cancer." *Genome* 51(12): 1032-9.
- Cantin, G. T., J. L. Stevens, et al. (2003). "Activation domain-mediator interactions promote transcription preinitiation complex assembly on promoter DNA." *Proc Natl Acad Sci U S A* 100(21): 12003-8.
- Caputi, M. and A. M. Zahler (2001). "Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family." *J Biol Chem* 276(47): 43850-9.
- Cartegni, L., S. L. Chew, et al. (2002). "Listening to silence and understanding nonsense: exonic mutations that affect splicing." *Nat Rev Genet* 3(4): 285-98.
- Cartegni, L., J. Wang, et al. (2003). "ESEfinder: A web resource to identify exonic splicing enhancers." *Nucleic Acids Res* 31(13): 3568-71.

- Cenini, G., R. Sultana, et al. (2008). "Elevated levels of pro-apoptotic p53 and its oxidative modification by the lipid peroxidation product, HNE, in brain from subjects with amnesic mild cognitive impairment and Alzheimer's disease." *J Cell Mol Med* 12(3): 987-94.
- Chabot, B., C. LeBel, et al. (2003). "Heterogeneous nuclear ribonucleoprotein particle A/B proteins and the control of alternative splicing of the mammalian heterogeneous nuclear ribonucleoprotein particle A1 pre-mRNA." *Prog Mol Subcell Biol* 31: 59-88.
- Chalkley, G. E. and C. P. Verrijzer (1999). "DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator." *EMBO J* 18(17): 4835-45.
- Chipuk, J. E. and D. R. Green (2006). "Dissecting p53-dependent apoptosis." *Cell Death Differ* 13(6): 994-1002.
- Chow, L. T., R. E. Gelinas, et al. (1977). "An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA." *Cell* 12(1): 1-8.
- Connell, C. M., S. P. Wheatley, et al. (2008). "Nuclear survivin abrogates multiple cell cycle checkpoints and enhances viral oncolysis." *Cancer Res* 68(19): 7923-31.
- Corden, J., B. Wasylyk, et al. (1980). "Promoter sequences of eukaryotic protein-coding genes." *Science* 209(4463): 1406-14.
- Cortes, P., O. Flores, et al. (1992). "Factors involved in specific transcription by mammalian RNA polymerase II: purification and analysis of transcription factor IIA and identification of transcription factor IIJ." *Mol Cell Biol* 12(1): 413-21.
- Cory, S. and J. M. Adams (2002). "The Bcl2 family: regulators of the cellular life-or-death switch." *Nat Rev Cancer* 2(9): 647-56.
- Cory, S., D. C. Huang, et al. (2003). "The Bcl-2 family: roles in cell survival and oncogenesis." *Oncogene* 22(53): 8590-607.
- Cowper, A. E., J. F. Caceres, et al. (2001). "Serine-arginine (SR) protein-like factors that antagonize authentic SR proteins and regulate alternative splicing." *J Biol Chem* 276(52): 48908-14.
- Davison, B. L., J. M. Egly, et al. (1983). "Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences." *Nature* 301(5902): 680-6.
- de Almeida, S. F., A. R. Grosso, et al. (2011). "Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36." *Nat Struct Mol Biol* 18(9): 977-83.
- Degterev, A., M. Boyce, et al. (2003). "A decade of caspases." *Oncogene* 22(53): 8543-67.

- Deng, W. and S. G. Roberts (2005). "A core promoter element downstream of the TATA box that is recognized by TFIIB." *Genes Dev* 19(20): 2418-23.
- Dikstein, R., S. Ruppert, et al. (1996). "TAFII250 is a bipartite protein kinase that phosphorylates the base transcription factor RAP74." *Cell* 84(5): 781-90.
- Dotson, M. R., C. X. Yuan, et al. (2000). "Structural organization of yeast and mammalian mediator complexes." *Proc Natl Acad Sci U S A* 97(26): 14307-10.
- Expert-Bezancon, A., J. P. Le Caer, et al. (2002). "Heterogeneous nuclear ribonucleoprotein (hnRNP) K is a component of an intronic splicing enhancer complex that activates the splicing of the alternative exon 6A from chicken beta-tropomyosin pre-mRNA." *J Biol Chem* 277(19): 16614-23.
- Faiger, H., M. Ivanchenko, et al. (2007). "Nearest-neighbor non-additivity versus long-range non-additivity in TATA-box structure and its implications for TBP-binding mechanism." *Nucleic Acids Res* 35(13): 4409-19.
- Fisette, J. F., J. Toutant, et al. (2010). "hnRNP A1 and hnRNP H can collaborate to modulate 5' splice site selection." *RNA* 16(1): 228-38.
- Fondell, J. D., H. Ge, et al. (1996). "Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex." *Proc Natl Acad Sci U S A* 93(16): 8329-33.
- Freiman, R. N., S. R. Albright, et al. (2002). "Redundant role of tissue-selective TAF(II)105 in B lymphocytes." *Mol Cell Biol* 22(18): 6564-72.
- Gangloff, Y. G., C. Romier, et al. (2001). "The histone fold is a key structural motif of transcription factor TFIID." *Trends Biochem Sci* 26(4): 250-7.
- Gangloff, Y. G., S. Werten, et al. (2000). "The human TFIID components TAF(II)135 and TAF(II)20 and the yeast SAGA components ADA1 and TAF(II)68 heterodimerize to form histone-like pairs." *Mol Cell Biol* 20(1): 340-51.
- Graveley, B. R. (2000). "Sorting out the complexity of SR protein functions." *RNA* 6(9): 1197-211.
- Han, S. P., Y. H. Tang, et al. (2010). "Functional diversity of the hnRNPs: past, present and perspectives." *Biochem J* 430(3): 379-92.
- Hoffmann, A. and R. G. Roeder (1996). "Cloning and characterization of human TAF20/15. Multiple interactions suggest a central role in TFIID complex formation." *J Biol Chem* 271(30): 18194-202.
- Hofmann, Y. and B. Wirth (2002). "hnRNP-G promotes exon 7 inclusion of survival motor neuron (SMN) via direct interaction with Htra2-beta1." *Hum Mol Genet* 11(17): 2037-49.

Holstege, F. C., P. C. van der Vliet, et al. (1996). "Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH." *EMBO J* 15(7): 1666-77.

House, A. E. and K. W. Lynch (2008). "Regulation of alternative splicing: more than just the ABCs." *J Biol Chem* 283(3): 1217-21.

Hsu, S. Y. and A. J. Hsueh (1998). "A splicing variant of the Bcl-2 member Bok with a truncated BH3 domain induces apoptosis but does not dimerize with antiapoptotic Bcl-2 proteins in vitro." *J Biol Chem* 273(46): 30139-46.

Jackson, I. J. (1991). "A reappraisal of non-consensus mRNA splice sites." *Nucleic Acids Res* 19(14): 3795-8.

Juven-Gershon, T. and J. T. Kadonaga (2010). "Regulation of gene expression via the core promoter and the basal transcriptional machinery." *Dev Biol* 339(2): 225-9.

Kar, A., K. Fushimi, et al. (2011). "RNA helicase p68 (DDX5) regulates tau exon 10 splicing by modulating a stem-loop structure at the 5' splice site." *Mol Cell Biol* 31(9): 1812-21.

Kastan, M. B., O. Onyekwere, et al. (1991). "Participation of p53 protein in the cellular response to DNA damage." *Cancer Res* 51(23 Pt 1): 6304-11.

Kayagaki, N., N. Yamaguchi, et al. (1999). "Involvement of TNF-related apoptosis-inducing ligand in human CD4+ T cell-mediated cytotoxicity." *J Immunol* 162(5): 2639-47.

Keren, H., G. Lev-Maor, et al. (2010). "Alternative splicing and evolution: diversification, exon definition and function." *Nat Rev Genet* 11(5): 345-55.

Kerr, J. F., A. H. Wyllie, et al. (1972). "Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics." *Br J Cancer* 26(4): 239-57.

Killeen, M. T. and J. F. Greenblatt (1992). "The general transcription factor RAP30 binds to RNA polymerase II and prevents it from binding nonspecifically to DNA." *Mol Cell Biol* 12(1): 30-7.

Kischkel, F. C., S. Hellbardt, et al. (1995). "Cytotoxicity-dependent APO-1 (Fas/CD95)-associated proteins form a death-inducing signaling complex (DISC) with the receptor." *EMBO J* 14(22): 5579-88.

Kischkel, F. C., D. A. Lawrence, et al. (2001). "Death receptor recruitment of endogenous caspase-10 and apoptosis initiation in the absence of caspase-8." *J Biol Chem* 276(49): 46639-46.

Klinck, R., A. Bramard, et al. (2008). "Multiple alternative splicing markers for ovarian cancer." *Cancer Res* 68(3): 657-63.

- Kolasinska-Zwierz, P., T. Down, et al. (2009). "Differential chromatin marking of introns and expressed exons by H3K36me3." *Nat Genet* 41(3): 376-81.
- Krummel, D. A., K. Nagai, et al. (2010). "Structure of spliceosomal ribonucleoproteins." *F1000 Biol Rep* 2.
- Kuribayashi, K. and W. S. El-Deiry (2008). "Regulation of programmed cell death by the p53 pathway." *Adv Exp Med Biol* 615: 201-21.
- Kutach, A. K. and J. T. Kadonaga (2000). "The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters." *Mol Cell Biol* 20(13): 4754-64.
- Ladd, A. N. and T. A. Cooper (2002). "Finding signals that regulate alternative splicing in the post-genomic era." *Genome Biol* 3(11): reviews0008.
- Lago, C., E. Clerici, et al. (2005). "The Arabidopsis TFIID factor AtTAF6 controls pollen tube growth." *Dev Biol* 285(1): 91-100.
- Lagrange, T., A. N. Kapanidis, et al. (1998). "New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB." *Genes Dev* 12(1): 34-44.
- Lavrik, I. N., A. Golks, et al. (2005). "Caspases: pharmacological manipulation of cell death." *J Clin Invest* 115(10): 2665-72.
- Le Hir, H., A. Nott, et al. (2003). "How introns influence and enhance eukaryotic gene expression." *Trends Biochem Sci* 28(4): 215-20.
- Leurent, C., S. Sanders, et al. (2002). "Mapping histone fold TAFs within yeast TFIID." *EMBO J* 21(13): 3424-33.
- Lewis, B. A., T. K. Kim, et al. (2000). "A downstream element in the human beta-globin promoter: evidence of extended sequence-specific transcription factor IID contacts." *Proc Natl Acad Sci U S A* 97(13): 7172-7.
- Li, H., H. Zhu, et al. (1998). "Cleavage of BID by caspase 8 mediates the mitochondrial damage in the Fas pathway of apoptosis." *Cell* 94(4): 491-501.
- Lim, C. Y., B. Santoso, et al. (2004). "The MTE, a new core promoter element for transcription by RNA polymerase II." *Genes Dev* 18(13): 1606-17.
- Long, J. C. and J. F. Cáceres (2009). "The SR protein family of splicing factors: master regulators of gene expression." *Biochem J* 417(1): 15-27.
- Luco, R. F., Q. Pan, et al. (2010). "Regulation of alternative splicing by histone modifications." *Science* 327(5968): 996-1000.

- Makeyev, A. V. and S. A. Liebhaber (2002). "The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms." *RNA* 8(3): 265-78.
- Makeyev, E. V., J. Zhang, et al. (2007). "The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing." *Mol Cell* 27(3): 435-48.
- Malik, S., W. Gu, et al. (2000). "The USA-derived transcriptional coactivator PC2 is a submodule of TRAP/SMCC and acts synergistically with other PCs." *Mol Cell* 5(4): 753-60.
- Maniatis, T. and R. Reed (2002). "An extensive network of coupling among gene expression machines." *Nature* 416(6880): 499-506.
- Marchand, V., M. Santerre, et al. (2011). "Identification of protein partners of the human immunodeficiency virus 1 tat/rev exon 3 leads to the discovery of a new HIV-1 splicing regulator, protein hnRNP K." *RNA Biol* 8(2): 325-42.
- Marshall, N. F., J. Peng, et al. (1996). "Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase." *J Biol Chem* 271(43): 27176-83.
- Martinez-Contreras, R., J. F. Fisette, et al. (2006). "Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing." *PLoS Biol* 4(2): e21.
- Moore, M. J., Q. Wang, et al. "An alternative splicing network links cell-cycle control to apoptosis." *Cell* 142(4): 625-36.
- Muller, F., M. A. Demeny, et al. (2007). "New problems in RNA polymerase II transcription initiation: matching the diversity of core promoters with a variety of promoter recognition factors." *J Biol Chem* 282(20): 14685-9.
- Muller, F. and L. Tora (2004). "The multicoloured world of promoter recognition complexes." *EMBO J* 23(1): 2-8.
- Muzio, M., B. R. Stockwell, et al. (1998). "An induced proximity model for caspase-8 activation." *J Biol Chem* 273(5): 2926-30.
- Newman, E. A., S. J. Muh, et al. (2006). "Identification of RNA-binding proteins that regulate FGFR2 splicing through the use of sensitive and specific dual color fluorescence minigene assays." *RNA* 12(6): 1129-41.
- Nikolov, D. B., H. Chen, et al. (1995). "Crystal structure of a TFIIB-TBP-TATA-element ternary complex." *Nature* 377(6545): 119-28.
- Nilsen, T. W. and B. R. Graveley (2010). "Expansion of the eukaryotic proteome by alternative splicing." *Nature* 463(7280): 457-63.
- Nudler, E., A. Goldfarb, et al. (1994). "Discontinuous mechanism of transcription elongation." *Science* 265(5173): 793-6.

- O'Shea-Greenfield, A. and S. T. Smale (1992). "Roles of TATA and initiator elements in determining the start site location and direction of RNA polymerase II transcription." *J Biol Chem* 267(9): 6450.
- Ohkuma, Y., S. Hashimoto, et al. (1995). "Analysis of the role of TFIIE in basal transcription and TFIIH-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIE- α ." *Mol Cell Biol* 15(9): 4856-66.
- Orphanides, G., T. Lagrange, et al. (1996). "The general transcription factors of RNA polymerase II." *Genes Dev* 10(21): 2657-83.
- Parent, A. and M. Bisailon (2006). "[Synergy between transcription and mRNA processing events]." *Med Sci (Paris)* 22(6-7): 626-32.
- Pointud, J. C., G. Mengus, et al. (2003). "The intracellular localisation of TAF7L, a paralogue of transcription factor TFIID subunit TAF7, is developmentally regulated during male germ-cell differentiation." *J Cell Sci* 116(Pt 9): 1847-58.
- Polunovsky, V. A., C. H. Wendt, et al. (1994). "Induction of endothelial cell apoptosis by TNF α : modulation by inhibitors of protein synthesis." *Exp Cell Res* 214(2): 584-94.
- Reese, J. C. (2003). "Basal transcription factors." *Curr Opin Genet Dev* 13(2): 114-8.
- Riedl, S. J., P. Fuentes-Prior, et al. (2001). "Structural basis for the activation of human procaspase-7." *Proc Natl Acad Sci U S A* 98(26): 14790-5.
- Roeder, R. G. (1996). "The role of general initiation factors in transcription by RNA polymerase II." *Trends Biochem Sci* 21(9): 327-35.
- Sato, S., C. Tomomori-Sato, et al. (2004). "A set of consensus mammalian mediator subunits identified by multidimensional protein identification technology." *Mol Cell* 14(5): 685-91.
- Scheer, E., F. Delbac, et al. (2012). "TFIID TAF6-TAF9 complex formation involves the HEAT repeat-containing C-terminal domain of TAF6 and is modulated by TAF5 protein." *J Biol Chem* 287(33): 27580-92.
- Shao, H., M. Revach, et al. (2005). "Core promoter binding by histone-like TAF complexes." *Mol Cell Biol* 25(1): 206-19.
- Simard, M. J. and B. Chabot (2000). "Control of hnRNP A1 alternative splicing: an intron element represses use of the common 3' splice site." *Mol Cell Biol* 20(19): 7353-62.
- Simard, M. J. and B. Chabot (2002). "SRp30c is a repressor of 3' splice site utilization." *Mol Cell Biol* 22(12): 4001-10.

- Singh, N. N., R. N. Singh, et al. (2007). "Modulating role of RNA structure in alternative splicing of a critical exon in the spinal muscular atrophy genes." *Nucleic Acids Res* 35(2): 371-89.
- Singh, S. and M. Dikshit (2007). "Apoptotic neuronal death in Parkinson's disease: involvement of nitric oxide." *Brain Res Rev* 54(2): 233-50.
- Smale, S. T. and D. Baltimore (1989). "The "initiator" as a transcription control element." *Cell* 57(1): 103-13.
- Spies, N., C. B. Nielsen, et al. (2009). "Biased chromatin signatures around polyadenylation sites and exons." *Mol Cell* 36(2): 245-54.
- Stennicke, H. R., J. M. Jurgensmeier, et al. (1998). "Pro-caspase-3 is a major physiologic target of caspase-8." *J Biol Chem* 273(42): 27084-90.
- Stephens, R. M. and T. D. Schneider (1992). "Features of spliceosome evolution and function inferred from an analysis of the information at human splice sites." *J Mol Biol* 228(4): 1124-36.
- Strasser, A., L. O'Connor, et al. (2000). "Apoptosis signaling." *Annu Rev Biochem* 69: 217-45.
- Subramanian, T., S. Vijayalingam, et al. (2007). "Evidence for involvement of BH3-only proapoptotic members in adenovirus-induced apoptosis." *J Virol* 81(19): 10486-95.
- Suda, T. and S. Nagata (1994). "Purification and characterization of the Fas-ligand that induces apoptosis." *J Exp Med* 179(3): 873-9.
- Thomas, M. C. and C. M. Chiang (2006). "The general transcription machinery and general cofactors." *Crit Rev Biochem Mol Biol* 41(3): 105-78.
- Thompson, C. B. (1995). "Apoptosis in the pathogenesis and treatment of disease." *Science* 267(5203): 1456-62.
- Thut, C. J., J. L. Chen, et al. (1995). "p53 transcriptional activation mediated by coactivators TAFII40 and TAFII60." *Science* 267(5194): 100-4.
- Venables, J. P., R. Klinck, et al. (2009). "Cancer-associated regulation of alternative splicing." *Nat Struct Mol Biol* 16(6): 670-6.
- Wagner, E. J. and M. A. Garcia-Blanco (2001). "Polypyrimidine tract binding protein antagonizes exon definition." *Mol Cell Biol* 21(10): 3281-8.
- Wan, D., Y. Gong, et al. (2004). "Large-scale cDNA transfection screening for genes related to cancer development and progression." *Proc Natl Acad Sci U S A* 101(44): 15724-9.

- Wang, W., R. Nahta, et al. (2004). "TAFII70 isoform-specific growth suppression correlates with its ability to complex with the GADD45a protein." *Mol Cancer Res* 2(8): 442-52.
- Warf, M. B. and J. A. Berglund (2010). "Role of RNA structure in regulating pre-mRNA splicing." *Trends Biochem Sci* 35(3): 169-78.
- Westover, K. D., D. A. Bushnell, et al. (2004). "Structural basis of transcription: separation of RNA from DNA by RNA polymerase II." *Science* 303(5660): 1014-6.
- Wilhelm, E., M. Kornete, et al. (2010). "TAF6delta orchestrates an apoptotic transcriptome profile and interacts functionally with p53." *BMC Mol Biol* 11: 10.
- Wilhelm, E., F. X. Pelay, et al. (2008). "TAF6delta controls apoptosis and gene expression in the absence of p53." *PLoS One* 3(7): e2721.
- Wright, K. J., M. T. Marr, 2nd, et al. (2006). "TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter." *Proc Natl Acad Sci U S A* 103(33): 12347-52.
- Wu, S. Y., T. Zhou, et al. (2003). "Human mediator enhances activator-facilitated recruitment of RNA polymerase II and promoter recognition by TATA-binding protein (TBP) independently of TBP-associated factors." *Mol Cell Biol* 23(17): 6229-42.
- Xie, X., T. Kokubo, et al. (1996). "Structural similarity between TAFs and the heterotetrameric core of the histone octamer." *Nature* 380(6572): 316-22.
- Xu, C., Y. Liu, et al. (2010). "Integrative analysis of DNA copy number and gene expression in metastatic oral squamous cell carcinoma identifies genes associated with poor survival." *Mol Cancer* 9: 143.
- Yin, X. M., Z. N. Oltvai, et al. (1994). "BH1 and BH2 domains of Bcl-2 are required for inhibition of apoptosis and heterodimerization with Bax." *Nature* 369(6478): 321-3.
- Zhang, H., K. Somasundaram, et al. (1998). "BRCA1 physically associates with p53 and stimulates its transcriptional activity." *Oncogene* 16(13): 1713-21.
- Zhang, J., X. Sun, et al. (1998). "Intron function in the nonsense-mediated decay of beta-globin mRNA: indications that pre-mRNA splicing in the nucleus can influence mRNA translation in the cytoplasm." *RNA* 4(7): 801-15.
- Zou, H., Y. Li, et al. (1999). "An APAF-1/cytochrome c multimeric complex is a functional apoptosome that activates procaspase-9." *J Biol Chem* 274(17): 11549-56.